



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2020

On disjointness properties of some parabolic flows

Kanigowski, Adam ; Lemańczyk, Mariusz ; Ulcigrai, Corinna

Abstract: The Ratner property, a quantitative form of divergence of nearby trajectories, is a central feature in the study of parabolic homogeneous flows. Discovered by Marina Ratner and used in her 1980th seminal works on horocycle flows, it pushed forward the disjointness theory of such systems. In this paper, exploiting a recent variation of the Ratner property, we prove new disjointness phenomena for smooth parabolic flows beyond the homogeneous world. In particular, we establish a general disjointness criterion based on the switchable Ratner property. We then apply this new criterion to study disjointness properties of smooth time changes of horocycle flows and smooth Arnol'd flows on T^2 , focusing in particular on disjointness of distinct flow rescalings. As a consequence, we answer a question by Marina Ratner on the Möbius orthogonality of time-changes of horocycle flows. In fact, we prove Möbius orthogonality for all smooth time-changes of horocycle flows and uniquely ergodic realizations of Arnol'd flows considered.

DOI: <https://doi.org/10.1007/s00222-019-00940-y>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-183922>

Journal Article

Submitted Version

Originally published at:

Kanigowski, Adam; Lemańczyk, Mariusz; Ulcigrai, Corinna (2020). On disjointness properties of some parabolic flows. *Inventiones Mathematicae*, 221(1):1-111.

DOI: <https://doi.org/10.1007/s00222-019-00940-y>

ON DISJOINTNESS PROPERTIES OF SOME PARABOLIC FLOWS

ADAM KANIGOWSKI, MARIUSZ LEMAŃCZYK, AND CORINNA ULCIGRAI

ABSTRACT. The Ratner property, a quantitative form of divergence of nearby trajectories, is a central feature in the study of parabolic homogeneous flows. Discovered by Marina Ratner and used in her 1980th seminal works on horocycle flows, it pushed forward the disjointness theory of such systems. In this paper, exploiting a recent variation of the Ratner property, we prove new disjointness phenomena for smooth parabolic flows beyond the homogeneous world. In particular, we establish a general disjointness criterion based on the *switchable* Ratner property. We then apply this new criterion to study disjointness properties of smooth time changes of horocycle flows and smooth Arnol'd flows on \mathbb{T}^2 , focusing in particular on disjointness of distinct flow *rescalings*. As a consequence, we answer a question by Marina Ratner on the Möbius orthogonality of time-changes of horocycle flows. In fact, we prove Möbius orthogonality for all smooth time-changes of horocycle flows and uniquely ergodic realizations of Arnol'd flows considered.

Dedicated to the memory of Marina Ratner

CONTENTS

1. Introduction	2
Disjointness for time-changes of horocycle flows	3
Disjointness properties of Arnol'd flows	5
Consequences for Möbius disjointness.	6
Outline of the paper	7
2. Preliminaries: definitions, notation and some basic facts	8
2.1. Flows, joinings and disjointness	8
2.2. Special flows	9
2.3. Time changes of flows	9
2.4. Irrational rotations, Denjoy-Koksma inequality	11
2.5. Ratner properties	12
3. A criterion for disjointness	13
3.1. Almost linear reparametrizations	14
3.2. The disjointness criterion	16
4. Disjointness for time changes of horocycle flows (proof of Theorem 1.1)	19
4.1. Preliminaries and shearing properties of time changes of horocycle flows	19
4.2. Deviations of ergodic averages estimates	20
4.3. Proof of Theorem 1.1	22
5. Disjointness criterion for special flows	26
6. Arnol'd flows and Birkhoff sums estimates	31
6.1. Definition of a class of Arnol'd flows	31
6.2. Denjoy-Koksma estimates	32
6.3. Estimates of Birkhoff sums at return times	32
6.4. Estimates on Birkhoff sums of f	34
6.5. Estimates on Birkhoff sums of f'	35
6.6. Birkhoff sums of f'' , f''' and f'''' .	43

6.7. A combinatorial lemma	46
7. Disjointness in Arnol'd flows (proof of Theorem 1.2)	46
7.1. Set up and beginning of the proof	47
7.2. Forward or backward shearing and concluding arguments	49
7.3. Slow shearing properties: proof of Lemma 7.2	51
7.4. Splitting of orbits: proof of Lemma 7.3	53
8. Disjointness of time changes of horocycle flows and Arnol'd flows (proof of Theorem 1.3)	60
9. Time changes of horocycle flows and Sarnak's conjecture (answer to M. Ratner's question)	64
Appendix A. Consequences of shearing for time changes of horocycle flows	66
Appendix B. Ergodic averages for horocycle flows	71
Appendix C. Strong MOMO and USIC properties	73
Acknowledgements	75
References	75

1. INTRODUCTION

In this paper we study ergodic properties of *parabolic* dynamical systems. Classical examples of parabolic systems are horocycle flows acting on unit tangent bundles of compact surfaces with constant negative curvature. In the 1980's, Marina Ratner established many spectacular rigidity phenomena for horocycle flows. In particular, Ratner classified invariant measures and joinings (for basic ergodic theory notions, see Section 2): both invariant measures and joinings have to be *algebraic*. The key phenomenon used in M. Ratner's theorems is the *Ratner property* (originally, in [30], it was called H-property, renamed as the R-property in [36]) which describes a special (polynomial) way of divergence of nearby trajectories. In [31], M. Ratner showed that the R-property persists for smooth *time changes* of horocycle flows. This, in [31], allowed her to show that (smooth) time changes of horocycle flows share with horocycle flows similar rigidity phenomena.

There is no formal definition for a system to be parabolic. In view of the above, it seems that the Ratner property is one of characteristics making a system parabolic. For a long time there were no known examples of systems with the Ratner property beyond horocycle flows and their (smooth) time changes. The situation has changed within the last 12 years. K. Frączek and the second author showed in [15] that there exists a class of flows on \mathbb{T}^2 (smooth flows with one singular point) for which a variant of the R-property holds. However, the flows in [15] are not (globally) smooth. The first class of globally smooth flows with a variation of the Ratner property beyond the horocyclic world was given by B. Fayad and the first author in [10]. The class considered in [10] consists of (some) smooth mixing flows on the torus \mathbb{T}^2 with one fixed point (these are sometimes known as Arnol'd and Kochergin flows). This triggered further investigations and the first and third authors and J. Kułaga-Przymus in [21] proved that the same variant of the Ratner property holds for smooth Arnol'd flows on every surface of genus $g > 1$. The variation of the Ratner property considered in [10] and [21] is the so called *switchable Ratner property* (or *SR-property* for short). For this variant (whose definition is recalled in Section 2.5), one only requires that the controlled form of divergence for nearby trajectories in Ratner's property holds *either* in the future *or* in the past, depending on the initial points. A recent result

of G. Forni and the first author [14] shows that the SR-property holds also in the class of smooth (non-trivial) time changes of constant type *Heisenberg nilflows*.

All the above mentioned variations of the R-property (which are detailed in Section 2.5) were defined in order to have the same strong dynamical consequences of the original R-property itself, hence we will sometimes call them simply *Ratner properties*. In particular, all Ratner properties, as the original R-property does, restrict the type of self-joinings that the flow can have (see Section 2.5) and allow to enhance mixing properties.

The Ratner property (or its variations) of a flow imposes some restrictions not only on the set of self-joinings but also on the set of its joinings with another (ergodic) flow. In particular, we can ask whether for two systems sharing the same Ratner property it is possible to classify joinings between them. The first result in this direction can be found in Ratner's work [31], where she shows that two flows (\bar{h}_t) and (\tilde{h}_t) given by two smooth time changes of a horocycle flow (h_t) are disjoint, i.e. the only joining between them is the product measure, whenever the cocycles corresponding to the time changes are not *jointly cohomologous* (see Section 2.3 for basic definitions).

In the present paper, we establish a general disjointness criterion based on the SR-property (Theorem 3.1), which we then use to prove disjointness properties of horocycle flows, their smooth time changes and some Arnol'd flows on \mathbb{T}^2 , see Theorems 1.1, 1.2 and 1.3. The statement (as well as the proof) of the criterion, which is rather technical, is given in Section 3. We now pass to a description of our main new disjointness results.

Disjointness for time-changes of horocycle flows. To state the first of our results, we need to recall some basic definitions. Let $G = SL(2, \mathbb{R})$ and Γ be a lattice in G . Let

$$h_t := \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, \quad t \in \mathbb{R},$$

and, by an abuse of notation, let also (h_t) denote the *horocycle flow* acting on $M := G/\Gamma$ (by left multiplication by h_t) considered with Haar measure μ . For $\tau \in C^1(M)$, $\tau > 0$, let (\tilde{h}_t^τ) denote the time change of (h_t) given by τ , i.e. for $x \in M$,

$$\tilde{h}_t^\tau(x) = h_{u(x,t)}x, \quad \text{where } \int_0^{u(x,t)} \tau(h_s x) ds = t.$$

Let us remark that while a time change seems to be the easiest form of perturbation of a flow (indeed, it preserves orbits), a time change of a horocycle flow typically breaks the homogeneity of the original flow and therefore can introduce new spectral and disjointness phenomena (see, for example, the results below). Thus, (smooth) time changes of horocycle flows constitute a fundamental class of non-homogeneous parabolic flows.

We will be particularly interested in *rescalings* of a flow (R_t) acting on a probability standard space (Z, \mathcal{D}, κ) . Given a real number $p > 0$, by the *p-rescaling* of (R_t) we simply mean the flow (R_{pt}) (in which the time was rescaled by the factor p). Notice that when p is an integer, the time-one map of p -rescaling coincides with the p -power of the time-one map R_1 of the flow, so considering rescalings is an analogous operation to considering the powers of a given transformation. In the case of a horocycle flow (or, more generally, a unipotent flow on a homogenous space), rescalings are also called *algebraic reparametrizations* since the time-changed flow is still of algebraic nature. We are interested in the situation when two rescalings (R_{pt}) and (R_{qt}) , where $p, q > 0$ and $p \neq q$, are disjoint, i.e. the only joining between them is product measure. Let us recall that the notion of disjointness was introduced by Furstenberg in [17] to study how *different* two systems can be. In particular,

if two flows are isomorphic, the isomorphism yields a non-trivial joining, so that the two flows cannot be disjoint (in fact, disjoint flows do not have a non-trivial common factor).

Notice that two different rescalings of a horocycle flow (h_t) are *never* disjoint. Indeed, if (g_s) denotes the *geodesic flow* given by left multiplication by

$$g_s := \begin{pmatrix} e^s & 0 \\ 0 & e^{-s} \end{pmatrix}$$

on $SL(2, \mathbb{R})/\Gamma$, the renormalization equation

$$(1.1) \quad h_t g_s = g_s h_{e^{-2s}t}, \quad \forall t, s \in \mathbb{R},$$

yields that, for any positive $p \neq q$, the flows (h_{pt}) and (h_{qt}) are conjugated by g_s with $s = -\frac{\log(q/p)}{2}$ (and hence are not disjoint).

However, for a large class of smooth time changes of a horocycle flow, we have the following result on disjointness of rescalings. In Theorem 1.2, we consider the class $B^+(M)$ of W^6 -smooth *positive* functions (time changes) which have non-trivial support outside of the discrete series (see Section 4.2 for the relevant definitions which require some basic notions from the representation theory of $SL(2, \mathbb{R})$).

Theorem 1.1. *Assume that Γ is cocompact. Assume moreover that the cocycle determined by $\tau \in B^+(M)$ is not a quasi-coboundary. Then for any positive $p, q \in \mathbb{R} \setminus \{0\}$, $p \neq q$, the flow rescalings (\tilde{h}_{pt}^τ) and (\tilde{h}_{qt}^τ) are disjoint.*

Notice that the above theorem does not hold if the cocycle given by τ is a quasi-coboundary, since in this case, (\tilde{h}_t^τ) is isomorphic to (h_t) and we have already remarked that the theorem fails for any horocycle flow because of the renormalization equation (1.1). Thus, non-trivial and sufficiently smooth time changes of (h_t) have *always* better disjointness properties than the horocycle flow itself.

Theorem 1.1 is proved using the new disjointness criterion given by Theorem 3.1 below, together with quantitative estimates on deviations of ergodic averages which follows from the works of Flaminio and Forni [12] and Bufetov and Forni [6] (it can be deduced from the results in [6], see Appendix B).

Let us remark that the question on possible disjointness of different rescalings of time changes of a horocycle flow is implicit in [31]. Indeed, from the results in [31], one can deduce (see [13] for details) that, for a smooth time change (\tilde{h}_t^v) of the horocycle flow, the rescalings (\tilde{h}_{pt}^v) and (\tilde{h}_{qt}^v) are disjoint if and only if the cocycles determined by v and $v \circ g_{-r}$, where $r := \frac{1}{2} \log(q/p)$, are not *jointly cohomologous* (see [31] for the definition).

Shortly after proving our result, and motivated by it, L. Flaminio and G. Forni informed us that an alternative proof of Theorem 1.1 can be obtained from the above mentioned disjointness result by Ratner [31], together with some considerations on the cohomological equation derived from their work in [12], and that, it also holds for an arbitrary $v \in W^6(M)$ such that the cocycle generated by v is not quasi-coboundary, see [13]. The main result proved in [13] is indeed that this absence of joint cohomology between the cocycles determined by v and $v \circ g_{-r}$ holds for $r \neq 0$ whenever the cocycle determined by v is not a quasi-coboundary. Note that the latter result is also *a posteriori* implied by our Theorem 1.1 when $v \in B^+(M)$.

We stress that our proof of Theorem 1.1 does *not* rely on Ratner's result from [31] and hence is perhaps more direct. It has furthermore the advantage of showing how deviations of ergodic averages and, in particular, quantitative *shearing phenomena* (through our disjointness criterion) play directly a role in disjointness of rescalings. While it might

be possible to adapt Flaminio and Forni's approach, which seem to rely essentially on representation theory only, to generalizations to other unipotent flows, our criterion of disjointness (Theorem 3.1), which requires as an input only dynamical properties of the flow, can be applied to study many other parabolic flows with similar quantitative shearing properties *outside* of the homogeneous world (see the further results in this paper on which we will detail shortly, as well as the conclusive remark at the end of Introduction mentioning other recent works based on our criterion, see [4, 8, 14]), thus providing a unifying approach to disjointness questions.

Disjointness properties of Arnol'd flows. Another important class of *parabolic flows* is given by *area-preserving flows on surfaces*. Let us define the class of *Arnol'd flows*, which, as we explain below, are flows which arise naturally when studying area preserving flows on a surface of genus one. A flow from this class has a special representation (see Section 2.2 for definitions) over an irrational rotation $R_\alpha(x) = x + \alpha$ on \mathbb{T} and under a roof function $f : \mathbb{T} \rightarrow \mathbb{R}_{>0}$ which (identifying \mathbb{T} with $[0, 1)$) has the form

$$f(x) = -A_- \log x - A_+ \log(1 - x) + h(x),$$

where $h \in C^5(\mathbb{T})$ and $A_- \neq A_+$ (both A_-, A_+ are positive numbers), see Figure 1. We will denote such flows by $((R_\alpha)_t^f)$.

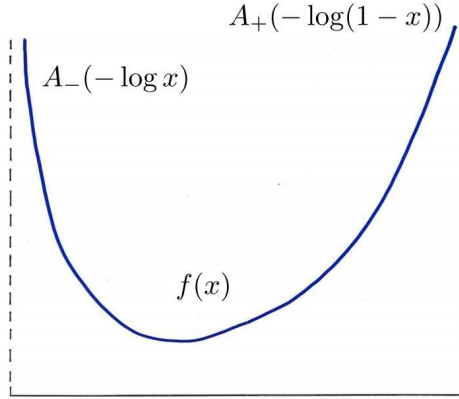


FIGURE 1. Special representation of an Arnol'd flow.

A motivation to study Arnol'd flows comes from considering smooth area-preserving flows on the torus \mathbb{T}^2 , with the simplest possible critical points, namely a center and a simple saddle (as in Figure 2). These flows are also known as *multi-valued* or *locally Hamiltonian* flows (see in particular the works by Novikov and its school [29]) since in suitable coordinates they are locally given by $\dot{x} = \frac{\partial H}{\partial y}$, $\dot{y} = -\frac{\partial H}{\partial x}$, where H is a local Hamiltonian. The minimal components of a *typical* such flow consist of invariant circles (which foliates a neighbourhood of the center, see Figure 2) and a non-trivial *minimal component* (supported on a torus minus a disk). The restriction (ϕ_t) of the flow to the minimal component admits a special representation of the form described above. All this was first remarked by Arnol'd in the paper [3] (from which the name *Arnol'd flows* comes), where mixing of (ϕ_t) was conjectured.

Our next result shows disjointness of rescalings of Arnol'd flows:

Theorem 1.2. *There exists a full Lebesgue measure set $\mathcal{D} \subset \mathbb{T}$ such that for every $\alpha \in \mathcal{D}$ and every $p, q > 0$, $\frac{p}{q} \notin \left\{1, \frac{A_-}{A_+}, \frac{A_+}{A_-}\right\}$, the flows $((R_\alpha)_{pt}^f)$ and $((R_\alpha)_{qt}^f)$ are disjoint.*

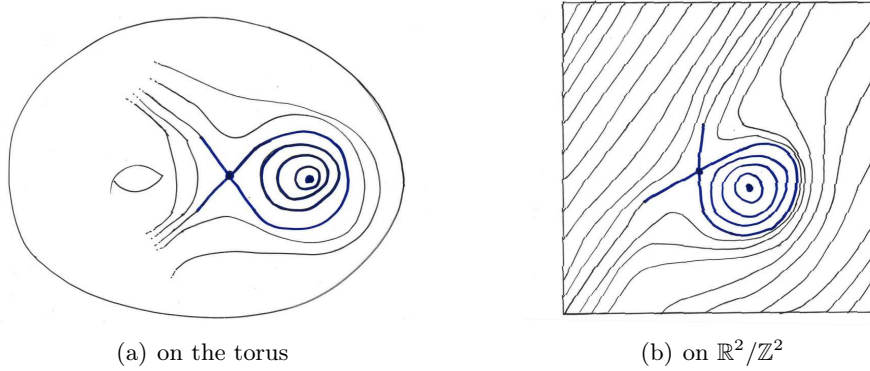


FIGURE 2. A locally Hamiltonian flow on \mathbb{T}^2 with a periodic orbits island and a minimal component (Arnol'd flow).

The proof of Theorem 1.2 which is done in Section 7 (and uses the results of Sections 6) is also based on the disjointness criterion given by Theorem 3.1.

Theorems 1.2 and 1.1 show an interesting phenomenon: the Ratner property in the non-homogeneous setting yields *stronger* dynamical consequences than in the homogeneous one. The reason (which will be explained in more details in Section 5) is that the Ratner property for a horocycle flow (hence in the homogeneous setting) depends only on the distance between the initial points and not on their position in the space. In all other examples of a Ratner property (in particular, for smooth non-trivial time changes of horocycle flows) divergence depends both on their distance and, what is more important, on their position in the space. This allows one to get stronger rigidity conclusions for the set of joinings.

Finally, our third main result, which is once again based on the disjointness criterion given by Theorem 3.1, shows that the two classes of parabolic flows considered so far, namely smooth time changes of horocycle flows and Arnol'd flows, are typically disjoint:

Theorem 1.3. *There exists a full Lebesgue measure set of rotation numbers $\mathcal{D} \subset \mathbb{T}$ such that for every $\alpha \in \mathcal{D}$, every horocycle flow (h_t) (with Γ cocompact) and every $\tau \in C^1(M)$, the Arnol'd flow $((R_\alpha)_t^f)$ and the time-change flow (\tilde{h}_t) are disjoint.*

The proof of Theorem 1.3 is at the heart of Section 8. Notice that since the result holds for *all* smooth time changes, it implies in particular disjointness between (homogeneous) horocycles flows and area-preserving flows in the class of Arnol'd flows considered in Theorem 1.2 (sets \mathcal{D} in both theorems are the same).

Consequences for Möbius disjointness. We now briefly discuss a relation between Theorems 1.1 and 1.2 with the *Möbius disjointness problem* [34] which is lately under intensive study, see e.g. survey [11]. If (T_t) is a continuous flow on a compact metric space X then it is called *Möbius disjoint* if

$$(1.2) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} F(T_{nt}x) \mu(n) = 0$$

for all $F \in C(X)$, $t \in \mathbb{R}$ and $x \in X$, where μ stands for the classical Möbius function (whose definition is recalled in Section 9). Sarnak conjectured in [34] that all zero entropy homeomorphisms are Möbius disjoint. As recent development shows, Sarnak's conjecture is very likely equivalent to the famous Chowla conjecture on correlations of the Möbius function, see [11] for details.

Möbius disjointness for horocycle flows has been proved by Bourgain, Sarnak and Ziegler in [5]. A few years ago M. Ratner asked about Möbius disjointness of (smooth) time changes of horocycle flows (see Question 7 in [11]). Since the flows appearing in Theorems 1.1 and 1.2 are totally ergodic (in fact, they are known to be mixing, see respectively [35] for Arnol'd flows and [28] for time changes of horocycle flows), in view of the so called Kátai-Bourgain-Sarnak-Ziegler criterion for orthogonality in [5] (recalled in Section 9), our results in particular imply that Möbius disjointness holds for all smooth time-changes in $B^+(M)$ of horocycle flows and for all uniquely ergodic models of Arnol'd flows (see Section 9 for more details). In particular, Theorem 1.1 answers positively Ratner's question on Möbius disjointness (Question 7 in [11]).

A question which has seen a recent surge of interest is *uniform convergence* in Möbius disjointness, namely whether the convergence in (1.2) is *uniform* in $x \in X$. It is for example shown in [2] that Sarnak's conjecture is equivalent to its uniform convergence form. While on the level of Möbius disjointness we cannot see any difference between horocycle flows and their (non-trivial) time changes, this situation changes drastically when we ask about uniformity. Indeed, Theorem 1.1 not only answers positively Ratner's question on Möbius disjointness, but it also yields uniform convergence (in $x \in M$) in (1.2). On the other hand, the question of whether uniform convergence holds for horocycle flows themselves remains largely open, see Question 5 in [11]. Thus, this is another instance, where in the non-homogeneous set up one can prove stronger properties than in the homogeneous world.

In the case of Arnol'd flows and the corresponding locally Hamiltonian flows, however, while we show Möbius disjointness of any uniquely ergodic realization, we have been unable until now to show uniformity of convergence in (1.2). Uniform convergence for Möbius disjointness, especially in its stronger form of so called convergence on (typical) short intervals, will be discussed in some detail in Section 9.

We conclude by saying that the (parabolic) disjointness criterion (Theorem 3.1) seems to be a general tool in studying disjointness of systems with a Ratner property. A variant of this criterion is used in a recent work of G. Forni and the first author [14], where disjointness properties of time changes of Heisenberg nilflows are studied and also in recent works of P. Berk and the first author [4] and C. Dong and the first author [8] for getting disjointness of some classes of interval exchange transformations (IET's) and translation flows.

Outline of the paper. The structure of the paper is as follows. In Section 2, we recall some basic definitions and properties, in particular about flows and special flows, joinings and irrational rotations on the circle. We also recall, for the convenience of the reader, the definition of the SR-property (a recent variation of the classical Ratner property), which is not directly used, but implicitly drives many of the exploited phenomena. In Section 3, we state and prove the new disjointness criterion (Theorem 3.1). Section 4 is devoted to the proof of Theorem 1.1 on disjointness for time changes of horocycle flows which gives the first illustration of the use of our criterion. In Section 5, we state a disjointness criterion in the language of special flows and we then show how it can be deduced from our main criterion on disjointness. The proof of the disjointness result for Arnol'd flows, i.e. Theorem 1.2, takes Sections 6 and 7: in Sections 6 we first prove some estimates on the growth of Birkhoff sums of derivatives of the roof function, then in Sections 7 we use them in order to quantify shearing and derive the splitting and slow drift of nearby orbits, properties which are needed to apply the disjointness criterion. In Section 8, we show that Arnol'd flows from the considered class are disjoint with smooth time changes of horocycle

flows (Theorem 1.3). Finally, in Section 9 we discuss consequences of our main results to the problem of Möbius disjointness. Three appendices follow. In the first two, we give the proofs of two technical results which are used in the proof of Theorem 1.1: in the first one, Appendix A, we prove some quantitative shearing properties for time-changes of horocycle flows; in Appendix B we state a result on ergodic averages of a special family of smooth functions which can be deduced from [6] (we thank G. Forni for the proof). Finally, in Appendix C, we present a missing link in the literature, namely, the equivalence between two different kinds of convergence on short intervals.

2. PRELIMINARIES: DEFINITIONS, NOTATION AND SOME BASIC FACTS

2.1. Flows, joinings and disjointness. Assume that (Z, \mathcal{D}, κ) is a probability standard Borel space. If Z is additionally a metric space (with a metric d), then we will also write $(Z, \mathcal{D}, \kappa, d)$ to emphasize the role of d . By $\text{Aut}(Z, \mathcal{D}, \kappa)$, we denote the set of all automorphisms of (Z, \mathcal{D}, κ) , i.e. $R \in \text{Aut}(Z, \mathcal{D}, \kappa)$ if $R : Z \rightarrow Z$ is a bi-measurable, measure-preserving κ -a.e. bijection. Each $R \in \text{Aut}(Z, \mathcal{D}, \kappa)$ determines a unitary operator, also denoted by R , on $L^2(Z, \mathcal{D}, \kappa)$ given by $Rf := f \circ R$ for $f \in L^2(Z, \mathcal{D}, \kappa)$. Endowed with the weak operator topology of unitary operators, $\text{Aut}(Z, \mathcal{D}, \kappa)$ becomes a Polish group.

We will be mainly deal with *flows*, i.e. with measurable, measure-preserving \mathbb{R} -actions $(R_t)_{t \in \mathbb{R}} \subset \text{Aut}(Z, \mathcal{D}, \kappa)$. Measurability means that the map $(z, t) \rightarrow R_t z$ is measurable (in fact, it is equivalent to saying that the unitary representation $t \mapsto R_t$ is continuous). If no confusion arises, we will abbreviate notation $(R_t)_{t \in \mathbb{R}}$ to R_t . By the *centralizer* $C(R_t)$ of the flow R_t , we mean the set of all $W \in \text{Aut}(Z, \mathcal{D}, \kappa)$ that commute with the flow: $WR_t = R_t W$ (for each $t \in \mathbb{R}$) (in fact, $C(R_t)$ is a closed subgroup of $\text{Aut}(Z, \mathcal{D}, \kappa)$). It is well-known (see e.g. [7]) that if a flow R_t is ergodic (i.e. its only invariant sets are either of zero or full measure) then for Lebesgue a.e. $s \in \mathbb{R}$, the time automorphism R_s is ergodic. Given $0 \neq r \in \mathbb{R}$, by R_{rt} we denote the flow $t \mapsto R_{rt}$.

Assume that T_t, S_t are flows on probability standard Borel spaces (X, \mathcal{B}, μ) and (Y, \mathcal{C}, ν) , respectively. A *joining* between T_t and S_t is a $T_t \times S_t$ -invariant (for each $t \in \mathbb{R}$) probability measure on $(X \times Y, \mathcal{B} \otimes \mathcal{C})$ with the projections μ and ν , respectively. By $J(T_t, S_t)$ we denote the set of joinings between the flows T_t and S_t . Following Furstenberg [17], we say that T_t and S_t are *disjoint*, and we write $T_t \perp S_t$, if product measure $\mu \otimes \nu$ is the only member of $J(T_t, S_t)$. Note that if $T_t \perp S_t$ then at least one of the two flows must be ergodic. If both flows are ergodic, then by $J^e(T_t, S_t)$ we denote the subset of *ergodic joinings*, that is, of those $\rho \in J(T_t, S_t)$ which make the flow $(T_t \times S_t)_{t \in \mathbb{R}} \subset \text{Aut}(X \times Y, \mathcal{B} \otimes \mathcal{C}, \rho)$ ergodic (this set is always nonempty, as the ergodic decomposition of any joining yields a.e. ergodic component also a joining). Note also that if W yields an isomorphism of T_t and S_t then the formula

$$\rho(B \times C) := \mu(B \cap W^{-1}C), \quad B \in \mathcal{B}, C \in \mathcal{C},$$

determines a member of $J(T_t, S_t)$ which is ergodic if T_t (hence S_t) is ergodic. Note also that, for each $r \neq 0$, we have $J(T_t, S_t) = J(T_{rt}, S_{rt})$ with the equality $J^e(T_t, S_t) = J^e(T_{rt}, S_{rt})$ whenever the flows T_t and S_t are ergodic.

Unless it is stated otherwise, flows are assumed to be *free* \mathbb{R} -actions, i.e. $t \mapsto T_t x$ is 1-1 for μ -a.e. $x \in X$. If T_t is ergodic and $S_t = Id$ (for each $t \in \mathbb{R}$, so S_t is not a free \mathbb{R} -action) then $T_t \perp S_t$. In fact, whenever an automorphism $R \in \text{Aut}(Z, \mathcal{D}, \kappa)$ is ergodic then R is disjoint from the identity (defined on an arbitrary probability, standard Borel space). We recall here that joinings for automorphisms are defined similarly as for flows – we replace

\mathbb{R} -actions by \mathbb{Z} -actions, in fact, joinings can be defined for actions of more general groups, see e.g. [19] for more details.

Remark 2.1. Although, in general, ergodic properties of a time automorphism and the whole flow are different (e.g. a time automorphism can have more factors, or can have a larger centralizer), there is no difference if we consider disjointness of time automorphisms. More precisely, given $r \neq 0$, the flows T_{rt} and S_{rt} are disjoint if and only if the time automorphisms T_r and S_r are disjoint. Indeed, the \mathbb{Z} -subaction $(T_{rn} \times S_{rn})$ is cocompact, so if $\rho \in J(T_r, S_r)$ then $\frac{1}{r} \int_0^r \rho \circ (T_s \times S_s) ds \in J(T_{rt}, S_{rt})$.

2.2. Special flows. Assume that $R \in \text{Aut}(Z, \mathcal{D}, \kappa)$ is an ergodic automorphism and let $f : Z \rightarrow \mathbb{R}$ belongs to $L^1(Z, \mathcal{D}, \kappa)$, $f > 0$ (κ -a.e.). For $n > 0$, set $f^{(n)}(z) = \sum_{j=0}^{n-1} f(R^j z)$. Complemented by $f^{(0)}(z) = 0$ and $f^{(-n)}(z) := -f^{(n)}(T^{-n}z)$, we obtain the cocycle identity

$$(2.1) \quad f^{(m+n)}(z) = f^{(m)}(z) + f^{(n)}(R^m z)$$

true for κ -a.e. $z \in Z$ and all $m, n \in \mathbb{Z}$. By a *special flow* $R^f = (R^f_t)_{t \in \mathbb{R}}$ with the *base* R and the *roof function* f , we mean the \mathbb{R} -action given by

$$(2.2) \quad R^f_t(z, r) = (R^n z, r + t - f^{(n)}(z)),$$

where $n \in \mathbb{Z}$ is the only number satisfying $f^{(n)}(z) \leq r + t < f^{(n+1)}(z)$. Then T^f_t is a flow defined on the probability standard Borel space $(Z^f, \mathcal{D}^f, \kappa^f)$, where the space

$$Z^f = \{(z, r) \in Z \times \mathbb{R} : z \in Z, 0 \leq r < f(z)\},$$

consists of the area in the plane under the graph of the function f , \mathcal{D}^f is the restriction of the product σ -algebra $\mathcal{D} \otimes \mathcal{B}(\mathbb{R})$ to Z^f , and κ^f is the restriction of product measure $\kappa \otimes \lambda_{\mathbb{R}}$ to \mathcal{D}^f , normalized by $\int_Z f d\kappa$. It is not hard to see that in view of the ergodicity of R , the special flow R^f_t is also ergodic.

Remark 2.2. One can observe that if $r \neq 0$ then the two special flows

$$(R^f_{r-1t})_{t \in \mathbb{R}} \text{ and } (R^{rf}_t)_{t \in \mathbb{R}} \text{ are isomorphic.}$$

Indeed, the map $Z^f \ni (x, s) \mapsto (x, rs) \in Z^{rf}$ is measure-preserving and equivariant.

2.3. Time changes of flows. Assume that R_t is a flow on (Z, \mathcal{D}, κ) and let $v \in L^1(Z, \mathcal{D}, \kappa)$ be a positive function. The function v determines a cocycle over R_t (which, by abusing the notation, we will denote again by v), given by the formula

$$(2.3) \quad v(t, x) := \int_0^t v(R_s x) ds.$$

Recall that the cocycle property means that $v(t_1 + t_2, x) = v(t_1, x) + v(t_2, R_{t_1} x)$ for κ -a.e. $x \in Z$ and all $t_1, t_2 \in \mathbb{R}$. One can then show (see for example [7]) that for a.e. $x \in Z$ and all $t \in \mathbb{R}$, there exists a unique $u = u(t, x)$ such that

$$(2.4) \quad \int_0^u v(R_s x) ds = t$$

(Note that if $t < 0$, then $u < 0$.) Now, we can define $\tilde{R}_t(x) := R_{u(t, x)}(x)$ and this is indeed an \mathbb{R} -action as the function $u = u(t, x)$ satisfies the cocycle identity: $u(t_1 + t_2, x) = u(t_1, x) + u(t_2, \tilde{R}_{t_1} x)$. The new flow \tilde{R}_t has the same orbits as the original flow (hence it is ergodic if R_t was), and it preserves the measure $\tilde{\kappa} \ll \kappa$, where $\frac{d\tilde{\kappa}}{d\kappa} = v / \int_Z v d\kappa$. We will also use notation \tilde{R}_t^v for the time change if it is not clear which function v is used.

We say that v is a *quasi-coboundary* if $v(t, x) = \xi(x) - \xi(R_t x) + t$ for a measurable $\xi : Z \rightarrow \mathbb{R}$. If v is a quasi-coboundary, then the flows R_t and \tilde{R}_t are isomorphic. More generally, if $w : Z \rightarrow \mathbb{R}^+$ is another time change of R_t and if

$$v(t, x) - w(t, x) = \xi(x) - \xi(R_t x)$$

(for κ -a.e. $x \in Z$ and all $t \in \mathbb{R}$) then the two time changes \tilde{R}_t^v and \tilde{R}_t^w are isomorphic, where the isomorphism is given by the map $x \mapsto \tilde{R}_{\xi(x)}^w x$ (in the special case of the trivial time change of R_t given by $w = 1$, for which the associated cocycle is $w(x, t) \equiv t$, this reduces to the previous definition of quasi-coboundary). Note that on Z we consider, in general, two different absolutely continuous measures.

Remark 2.3. One can show that any special flow R_t^f , where $R \in \text{Aut}(Z, \mathcal{D}, \kappa)$, can be obtained from a time change of the *suspension* of R , i.e. from a time change of the special flow over R built under the constant function 1.

Remark 2.4. If R_t is a smooth flow on a manifold Z and v is also smooth then the time change flow \tilde{R}_t^v is a smooth flow. In fact, when $A : Z \rightarrow TZ$ is a vector field and

$$\frac{d}{dt} R_t x = A(R_t x)$$

and if $v : Z \rightarrow \mathbb{R}^+$ is smooth, then consider the vector field $\frac{1}{v}A$. There is a unique (smooth) flow \tilde{R}_t given by

$$\frac{d}{dt} \tilde{R}_t x = \frac{1}{v(\tilde{R}_t x)} A(\tilde{R}_t x)$$

and one can show that the flow \tilde{R}_t thus defined (which is clearly smooth by construction) coincides with the flow \tilde{R}_t^v defined at the beginning of the section.

Let us conclude this section with two very simple lemmas on time changes which will be helpful later.

Lemma 2.1. *Assume that Z is a compact metric space and $R_t : Z \rightarrow Z$ is a uniquely ergodic flow. Then, for any $v \in C(Z)$ a positive function with $\int_Z v d\kappa = 1$, we have*

$$\lim_{t \rightarrow \infty} \frac{t}{u(t, z)} = 1$$

uniformly in $z \in Z$.

Proof. Remark that for each (and, in fact, uniformly in) $z \in Z$, we have

$$\frac{t}{u(t, z)} = \frac{1}{u(t, z)} \int_0^{u(t, z)} v(R_s z) ds \xrightarrow{t \rightarrow \infty} \int_Z v d\kappa = 1.$$

The conclusion follows immediately. \square

The following shows that the study of rescalings of a flow can be reduced to the study of time changes.

Lemma 2.2. *Consider a flow $R_t : Z \rightarrow Z$ on (Z, \mathcal{D}, κ) , $v : Z \rightarrow \mathbb{R}^+$ integrable and let \tilde{R}_t^v be the corresponding time change. Fix $0 < p \in \mathbb{R}$ and consider \tilde{R}_t^{pv} . Then, for κ -a.e. $x \in Z$ and all $t \in \mathbb{R}$, we have*

$$\tilde{R}_t^{pv}(x) = \tilde{R}_{(1/p)t}^v(x).$$

Proof. For κ -a.e. $x \in Z$ and all $t \in \mathbb{R}$ we have $\int_0^{u(x, t)} (pv)(R_s x) ds = pt$. This can be interpreted by saying that $\tilde{R}_{pt}^{pv}(x) = R_{u(x, t)}(x)$. Hence, since $R_{u(x, t)}(x) = \tilde{R}_t(x)$, it gives the conclusion. \square

2.4. Irrational rotations, Denjoy-Koksma inequality. For each $x \in \mathbb{R}$, we set

$$\|x\| := \min(\{x\}, 1 - \{x\}), \quad \text{where } \{x\} = x - [x]$$

stands for the fractional part of x . By $\mathbb{T} = [0, 1)$ we denote the additive circle. Then, clearly, $\|\cdot\|$ determines a translation invariant metric on \mathbb{T} . By $\lambda_{\mathbb{T}}$ (or λ) we will denote Lebesgue measure on \mathbb{T} .

Let

$$[0, 1) \ni \alpha = [0; a_1, a_2, \dots] = \frac{1}{a_1 + \frac{1}{a_2 + \dots}}$$

be irrational. Then the positive integers a_i are called *partial quotients* of α . By setting

$$p_0 = 0, p_1 = 1, p_{n+1} = a_{n+1}p_n + p_{n-1}, \quad q_0 = 1, q_1 = a_1, q_{n+1} = a_{n+1}q_n + q_{n-1}$$

for $n \geq 2$, we obtain the sequences $(p_n), (q_n)$ of *numerators* and *denominators* of α , respectively. Then (see e.g. [23])

$$\frac{1}{2q_n q_{n+1}} < \left| \alpha - \frac{p_n}{q_n} \right| < \frac{1}{q_n q_{n+1}};$$

equivalently,

$$(2.5) \quad \frac{1}{2q_{n+1}} < \|q_n \alpha\| < \frac{1}{q_{n+1}}.$$

From this inequality, we can deduce the following elementary *spacing* properties of orbits which we will use several times later.

Lemma 2.3 (Spacing of orbits up to return times). *For any $x \in \mathbb{T}$ and $n \in \mathbb{N}$, the orbit $\{x + i\alpha, 0 \leq i < q_n\}$ is such that*

$$(2.6) \quad \|x + i\alpha - (x + j\alpha)\| \geq \frac{1}{2q_n}, \quad \text{for all } 0 \leq i \neq j < q_n.$$

On the other hand, in each interval of length $2/q_n$ there must be a point of the form $x + j\alpha$ for some $0 \leq j < q_n$.

Proof. For the first part, note that for $0 \leq i, j < q_n$, $\|x + i\alpha - (x + j\alpha)\| = \|(i - j)\alpha\|$ with $|i - j| < q_n$. Since $\frac{1}{2q_n} < \|q_{n-1}\alpha\|$ and $\|r\alpha\| \geq \|q_{n-1}\alpha\|$ whenever $|r| < q_n$, we have (2.6). For the second, assuming that $\alpha > \frac{p_n}{q_n}$, we have $k\alpha - \frac{kp_n}{q_n} < \frac{1}{q_{n+1}}$ for $0 \leq k < q_n$, so that in particular there exist $0 \leq i \neq j < q_n$ such that

$$\|x + i\alpha - (x + j\alpha)\| < \frac{1}{q_n} + \frac{1}{q_{n+1}} < \frac{2}{q_n}.$$

Treating the case $\alpha < \frac{p_n}{q_n}$ similarly, the second part follows. \square

If $f : \mathbb{T} \rightarrow \mathbb{R}$ has bounded variation then, for each $n \geq 0$ and $x \in \mathbb{T}$, we have the following *Denjoy-Koksma inequality*:

$$(2.7) \quad \left| f^{(q_n)}(x) - q_n \int_{\mathbb{T}} f d\lambda \right| \leq 2\text{Var}(f),$$

see e.g. [25].

We will also make use of so called *Ostrowski expansion* of a natural number N . Namely, we can represent N as

$$N = \sum_{j=1}^K b_j q_j,$$

where $0 \leq b_j \leq a_j$ (and $b_K \neq 0$) and if, for some j , we have $b_j = a_j$ then necessarily $b_{j+1} = 0$. Such a decomposition is unique, see e.g. [25].

2.5. Ratner properties. Our main disjointness criterion (Theorem 3.1) is based on quantitative divergence properties and is inspired by the Ratner property (originally defined in [30]) and in particular its recent variation known as the *switchable Ratner property* (SR-property), first introduced in [10]. While the SR-property is not explicitly used in the paper, it is implicitly present, both in formulation of the disjointness criterion and the phenomena driving it. We hence include in this section its formal definition followed by some explanations, as we think that the reader who is not familiar with it, might benefit by reading it before the following section.

Let (X, d) be a σ -compact metric space, \mathcal{B} the σ -algebra of Borel subsets of X and μ a Borel probability measure on X . Let $(T_t)_{t \in \mathbb{R}}$ be an ergodic flow acting on (X, \mathcal{B}, μ) . The Ratner property encodes a quantitative property of controlled divergence of nearby trajectories in the flow direction. Heuristically, we want that for *most* pairs of nearby points x, x' , the orbits of x, x' *split* in the flow direction (say at time $t = M$) by a definite amount (the *shift* p , which belongs to a fixed compact set P) and then *realign*, so that now $T_t(x)$ and the time-shifted orbit $T_{t+p}(x')$ are close, and stay close for a fixed proportion κM of the time it took to see the shift, namely for *most* times $t \in [M, M + L]$ where $L > \kappa M$ (this is made precise in Definition 2.1). One can see that this type of phenomenon is possible only for parabolic systems, in which orbits of nearby points diverge with polynomial or subpolynomial speed.

In the *switchable* variation, this phenomenon might not happen for positive times t (i.e. in the *future*), but only in the *past*, i.e. for $t < 0$ (and one can *switch* between exploiting the past or the future according to the pair of points x, x'). This is encoded by the following formal definition. (We comment below on the relation with the definitions in the literature, see Remark 2.5).

Definition 2.1 (SR-property). We say that the flow (T_t) has the *SR-property* (with *shifts set* P) if there exists a compact set $P \subset \mathbb{R} \setminus \{0\}$ (the *shifts set*) such that:

for every $\varepsilon > 0$ and $N \in \mathbb{N}$ there exist $\kappa = \kappa(\varepsilon)$, $\delta = \delta(\varepsilon, N)$ and a set $Z = Z(\varepsilon, N)$ with $\mu(Z) > 1 - \varepsilon$, such that:

for every $x, y \in Z$ with $d(x, y) < \delta$ and x not in the orbit of y , there exist $p = p(x, y) \in P$ and $M = M(x, y)$, $L = L(x, y)$ such that $M \geq N$ and $L \geq \kappa M$, and at least one of the following holds:

- (i) $d(T_t(x), T_{t+p}(y)) < \varepsilon$ for $t \in U \subset [M, M + L]$, or
- (ii) $d(T_t(x), T_{t+p}(y)) < \varepsilon$ for $-t \in U \subset [M, M + L]$,

where $U \subset [M, M + L]$ is such that its Lebesgue measure $|U|$ satisfies $|U| > (1 - \varepsilon)L$.

Referring to the heuristic explanation before the definition, *most* initial points is formalized by the set Z of arbitrarily large measure. The splitting and realignment with shift of nearby orbits should happen for arbitrarily large times (i.e. for every $N \geq 0$ there must be an $M \geq N$). Finally, (i) describes controlled splitting in the *future*, while (ii) describes controlled splitting in the *past*.

Remark 2.5. In the classical Ratner property, $P = \{1, -1\}$ and for all $x, y \in Z$ (i) has to be satisfied. The first extensions of the Ratner property were introduced by K. Frączek and M. Lemańczyk in [15] and [16] (as mentioned in the introduction) and amounted to allow P to be any finite set (the *finite Ratner property*) or any compact set $P \subset \mathbb{R} \setminus \{0\}$ (the *weak Ratner property*)¹. The *switchable* variation (where the possibility (ii) of controlled

¹The definition given in [15] and [16] is slightly different, in that the property is stated in terms of the time t maps of the flow, and the flow is said to have the corresponding Ratner properties if the set of times

splitting in the past is also allowed) was introduced by B. Fayad and the first author in [10].

All the variants of the Ratner properties are defined so that, if (T_t) has the SR-property, then it has the *finite extension of joinings property* (shortened as FEJ property, see [10]), which is a rigidity property that restricts the type of self-joinings that (T_t) can have (see [33, 15]).

3. A CRITERION FOR DISJOINTNESS

The main result of this section is the new disjointness criterion which the main results of the rest of the paper are based on. Since this criterion and the explanations in this section are inspired by the SR-property (even if the property is not used directly), the reader who is not familiar with it might benefit from reading Section 2.5 first.

The statement of the criterion is given in Section 3.2. Since it is rather technical, let us first explain the guiding ideas behind it. The criterion was devised and formulated so that it can be applied to prove disjointness of two flows which both have the SR-property (see Definition 2.1), so that in both flows one can observe a controlled form of divergence of nearby trajectories (for example polynomial divergence), but the speed of divergence for the two flows is different (for example for one flow is it linear, in the other quadratic).

The key *shearing phenomenon* exploited in the criterion is that, for pairs of two nearby points in the first system and two nearby points in the second, after some time (depending on both pairs of points) we will see a *relative divergence*, i. e. in one pair we will see a realignment with some shift $C_1 > 0$ in the flow direction, while in the other we will see a realignment with a shift $C_2 > 0$ with $C_1 \neq C_2$. To see that this is the case when both flows have the Ratner property (*not* switchable for now) but with different speeds, one can reason as follows. Using the Ratner or SR-property, one can find times $M_1 > 0$ and $M_2 > 0$ such that the first pair splits exactly by 1 at time M_1 and the second splits by exactly 1 at time M_2 . If the second pair has not yet not split at time M_1 (or symmetrically, first pair has not yet split at time M_2), then we are done with $C_1 = 1$ for one flow and $C_2 = 0$ for the other. The interesting case is when M_1 and M_2 are very close (assume for simplicity that $M_1 = M_2$). In this case one can use the fact that the speed of divergence is different to conclude that after some time, say $10M_1$, the two pairs will have split by a different amount (in the linear and quadratic case, it is a simple consequence of the fact that a line and parabola can have at most two points of intersection -at time M_1 - and after M_1 they start slowly diverging). This explains how the relative shearing appears in this situation.

Let us remark that if both flows have only the *switchable* Ratner property, one might have pairs of points for which one see the Ratner form of shearing only in the past for one flow, and only in the future for the other. Therefore, there is no hope to implement the heuristic described above on the *whole* (or large parts) of the space. An essential feature of the criterion is that for one of the two flows one can consider only pairs of points in *small* parts of space (whose measure is bounded from below, but not necessarily close to one), on which one sees controlled shearing *both* forward *and* backward. Thus, one can couple these pairs with pairs in the other flow which have the Ratner property in the future *or* the past to get the *relative divergence* explained above. We will come back to this point after the statement of the criterion in Section 3.2.

t such that T_t has the corresponding property is uncountable. This is clearly implied by the definition given here.

In order to give the precise formulation of the criterion, we now first need to give the definition of *almost linear* reparametrization.

3.1. Almost linear reparametrizations. In the disjointness criterion stated in the next Section 3.2, we will consider a class of time reparametrizations given by functions $t \mapsto a(t)$ which are *almost linear* (in the sense of Definition 3.1). Heuristically, we want the function $a(\cdot)$ to be "close" to a linear function, either being close to a piecewise linear function, which on each interval of continuity has the form $a(t) = t + R_i$ (we call this case (PAL) for *Piecewise Almost Linear*), or having derivative $a'(t)$ close to 1 (a case dabled (SAL), for *Smoothly Almost Linear*).

For a (measurable) set $I \subset \mathbb{R}$, $|I|$ denotes its Lebesgue measure.

Definition 3.1 (almost linear functions and good triples). Assume that $I = [u_1, u_2]$ is an interval, $a : I \rightarrow \mathbb{R}$, $U \subset I$ and $0 < d, \xi < 1$. We say that a is *almost linear* and that the triple (a, U, d) is ξ -good if at least one of the following holds:

(PAL) (*Piecewise Almost Linear function*) we can write U as $U = \bigcup_{i=1}^v (c_i, d_i)$, where $v \leq |I|/d$ and $|U| > (1 - \xi)|I|$, and for each $i = 1, \dots, v$, for $x \in (c_i, d_i)$ we have that $a(t) = t + R_i$ for some $R_i \in \mathbb{R}$ such that

$$R_1 \leq \xi|I|, \quad |R_{i+1} - R_i| < \xi \text{ for every } i = 1, \dots, v-1.$$

In this case we say that a is a *piecewise almost linear* function.

(SAL) (*Smooth Almost Linear function*) the function $a \in C^1(I)$, $U = I = [u_1, u_2]$ and we have that $|a(u_i) - u_i| \leq \xi|I|$ for $i = 1, 2$ and

$$1 - \xi < a'(t) \leq 1 + \xi, \quad \text{for every } t \in I.$$

In this case we say that a is a *smooth almost linear* function.

Remark 3.1. The notion of (PAL) for $a(\cdot)$ is introduced in order to deal with special flows: in this case, we need to take care of the fact that special flows are discontinuous close to the roof. This phenomenon produces the decomposition (c_i, d_i) , $i = 1, \dots, v$ in (PAL).

This definition of almost linear functions and good triples is given so that the following result (which is essentially based on a change of variables argument) holds. The lemma guarantees that if $a(\cdot)$ is (PAL) or (SAL) then the orbital integrals $\int_M^{M+L} f(T_t x) dt$ and $\int_M^{M+L} f(T_{a(t)} x) dt$ are close.

Lemma 3.1 (Closeness of almost linear good reparametrizations). Let $0 < d, \xi < 1$, $a : [M, M+L] \rightarrow \mathbb{R}$ and let $U \subset [M, M+L]$ be such that a is almost linear and (a, U, d) is ξ -good. Assume that $f : \mathbb{R} \rightarrow (0, 1]$ is measurable. Then

$$\frac{1}{L} \int_M^{M+L} f(a(t)) dt < \frac{1}{L} \int_M^{M+L} f(s) ds + 9\xi/d.$$

The rest of this subsection will be taken by the proof of this lemma.

Proof. Assume first that a is smoothly almost linear, that is, (SAL) is satisfied. Then

$$\left| \frac{1}{L} \int_M^{M+L} f(a(t)) dt - \frac{1}{L} \int_M^{M+L} f(a(t)) a'(t) dt \right| \leq \xi.$$

On the other hand, substituting $s = a(t)$, by the assumptions on the endpoints given by (SAL), we obtain that

$$\frac{1}{L} \int_M^{M+L} f(a(t)) a'(t) dt = \frac{1}{L} \int_{a(M)}^{a(M+L)} f(s) ds = \frac{1}{L} \int_M^{M+L} f(s) ds + 2\xi.$$

Combining the estimates, this concludes the proof in this case (since $d < 1$ and hence $3\xi < 9\xi/d$).

Assume now that a is piecewise almost linear, so (PAL) is satisfied. Let $A_s := \{1 \leq i \leq v : d_i - c_i \leq \xi\}$ (where s stays for *short* intervals). Let $U_s := \bigcup_{i \in A_s} [c_i, d_i]$ and $U_l := U \setminus U_s$ (where l stays for *large*). We then have

$$\begin{aligned} \left| \frac{1}{L} \int_M^{M+L} f(a(t)) dt - \frac{1}{L} \int_{U_l} f(a(t)) dt \right| &= \frac{1}{L} \int_{(I \setminus U) \cup U_s} f(a(t)) dt \\ &\leq \frac{|I \setminus U|}{L} + \frac{\xi v}{L} \leq \frac{\xi L}{L} + \frac{\xi L}{L d} < \frac{2\xi}{d} \end{aligned}$$

(where in the last inequality we used that $d < 1$). Moreover, substituting $r = a(t)$ on each (c_i, d_i) for $i \notin A_s$ and denoting by $a(x^\pm)$ respectively the right and left limit of the function $a(t)$ as $t \rightarrow x^\pm$, we get

$$\begin{aligned} \frac{1}{L} \int_{U_l} f(a(t)) dt &= \frac{1}{L} \sum_{\substack{i=1, \dots, v \\ i \notin A_s}} \int_{a(c_i^+)}^{a(d_i^-)} f(r) dr \\ (3.1) \quad &= \frac{1}{L} \sum_{\substack{i=1, \dots, v \\ i \notin A_s}} \int_{a(c_i^+)}^{a(d_i^-) - \xi} f(r) dr + \frac{1}{L} \sum_{\substack{i=1, \dots, v \\ i \notin A_s}} \int_{a(d_i^-) - \xi}^{a(d_i^-)} f(r) dr. \end{aligned}$$

Since $|A_s| \leq v \leq L/d$ (where $|A_s|$ denotes the cardinality of A_s), we have

$$\frac{1}{L} \sum_{\substack{i=1, \dots, v \\ i \notin A_s}} \int_{a(d_i^-) - \xi}^{a(d_i^-)} f(r) dr \leq \frac{\xi |A_s|}{L} \leq \frac{\xi}{d}.$$

Now, we estimate the first term in RHS of (3.1). We claim that the intervals $[a(c_i^+), a(d_i^-) - \xi]$ for $i \notin A_s$ are non-empty (since $i \notin A_s$ implies that $d_i - c_i > \xi$ and hence $a(d_i^-) - a(c_i^+) > \xi$) and *pairwise disjoint*. To see this, notice that by the definition (PAL) of almost linear function, for $i \notin A_s$, we have

$$a(c_i^+) = c_i + R_i \leq d_i + R_i - \xi = a(d_i^-) - \xi = d_i + R_i - \xi \leq c_{i+1} + R_{i+1} = a(c_{i+1}^+),$$

which proves that the intervals $[a(c_i^+), a(d_i^-) - \xi]$ are pairwise disjoint (and in increasing order). This now gives that

$$\frac{1}{L} \sum_{i=1, i \notin A_s}^v \int_{a(c_i^+)}^{a(d_i^-) - \xi} f(r) dr \leq \frac{1}{L} \int_{a(c_1^+)}^{a(d_v^-)} f(r) dr.$$

Remark now that, by the assumptions in (PAL) and recalling that $d < 1$, we have that

$$|R_v| = \left| \sum_{i=1}^{v-1} (R_{i+1} - R_i) + R_1 \right| \leq \xi v + \xi L \leq 2\xi \frac{L}{d}.$$

Hence, using this bound on R_v and the assumptions in (PAL) (and $d < 1$), we have that

$$\begin{aligned} \left| \frac{1}{L} \int_{a(c_1^+)}^{a(d_v^-)} f(r) dr - \frac{1}{L} \int_M^{M+L} f(r) dr \right| &\leq \frac{|a(c_1^+) - M| + |a(d_v^-) - M - L|}{L} \\ &\leq \frac{|c_1 - M| + |d_v - M - L| + |R_1| + |R_v|}{L} \leq 3\xi + 2\frac{\xi}{d}. \end{aligned}$$

Combining all the estimates together, we get the desired conclusion also in this case. This finishes the proof. \square

3.2. The disjointness criterion. Recall that we are considering measurable, measure-preserving \mathbb{R} -actions (i.e. flows) on probability standard Borel spaces. In fact, we will be constantly assuming that our configuration spaces are σ -compact metric spaces (considered with the Borel σ -algebras and probability Borel measures).

Let (T_t) and (S_t) be two weakly mixing flows acting on $(X, \mathcal{B}, \mu, d_1)$ and $(Y, \mathcal{C}, \nu, d_2)$, respectively. Given $A \subset X$, we set $V_\epsilon^1(A) := \{x \in X : d_1(x, A) < \epsilon\}$, and a similar notation $V_\epsilon^2(A')$ is used for a subset A' of Y . Let $P = \{p, -p\}$, $p \neq 0$.

Let us first state the criterion, then make some comments that connect it to the heuristics presented at the beginning of the section.

Theorem 3.1 (Disjointness criterion). *Let $0 < c < 1$. Assume that we have a sequence of sets*

$$(X_k) \subset \mathcal{B}, \quad \mu(X_k) \rightarrow \mu(X),$$

together with a sequence of automorphisms

$$(A_k) \subset \text{Aut}(X_k, \mathcal{B}|_{X_k}, \mu|_{X_k}), \quad \text{such that } A_k \rightarrow \text{Id uniformly}^2.$$

Assume moreover that for every $0 < \epsilon < 1$ and $N \in \mathbb{N}$ there exist a sequence

$$(E_k) = (E_k(\epsilon)) \subset \mathcal{B}, \quad \mu(E_k) \geq c\mu(X),$$

and $0 < \kappa = \kappa(\epsilon) < \epsilon$, $\delta = \delta(\epsilon, N) > 0$ and a set

$$Z = Z(\epsilon, N) \subset Y, \quad \nu(Z) \geq (1 - \epsilon)\nu(Y)$$

such that for all $y, y' \in Z$ satisfying $d_2(y, y') < \delta$, every k such that $d_1(A_k, \text{Id}) < \delta$ and every $x \in E_k \cap X_k$, $x' := A_k x$ there are

$$M \geq N, \quad L \geq 1, \quad \frac{L}{M} \geq \kappa, \quad \text{and } p \in P$$

for which at least one of the following holds:

$$(3.2) \quad \max(d_1(T_t x, T_{a(t)+p} x'), d_2(S_t y, S_{a(t)} y')) < \epsilon \text{ for } t \in U \subset [M, M+L]$$

or

$$(3.3) \quad \max(d_1(T_t x, T_{a(t)+p} x'), d_2(S_t y, S_{a(t)} y')) < \epsilon \text{ for } -t \in U \subset [M, M+L],$$

where $a := a_{x, x', y, y'} : [M, M+L] \rightarrow \mathbb{R}$ is extended by $a(-t) = a(t)$, and (a, U, c) is ϵ -good.

Then, the flows (T_t) and (S_t) are disjoint.

As we explained at the beginning of this section, the criterion is meant to be applied to two flows having the SR-property (or the Ratner property) but with different speed. The parameters (X_k, A_k, E_k) allows us to relax the SR-property for one of the flows: we only need to control a positive proportion of space (the sets E_k) and only in a favorite direction (that we control well, for instance, the geodesic direction $A_k = g_{1/k}$ for the horocycle

²This means that for each $\delta' > 0$, $d_1(A_k x, x) < \delta'$ for all k large enough and $x \in X_k$.

flow), we do not need to control *all* nearby points. The set Z is just the set, where the SR-property holds for the other flow. The almost linear function $a(t)$ describes the relative drift between points x, x' (or y, y') (for example, it would be $t + td(x, x')$ if the divergence is linear). Formulas (3.2) and (3.3) then describe the *relative shearing*, *either* in the future *or* in the past.

The rest of this section is devoted to the proof of the disjointness criterion.

Proof. Let $\rho \in J((T_t)_{t \in \mathbb{R}}, (S_t)_{t \in \mathbb{R}})$ be an ergodic joining, $\rho \neq \mu \times \nu$. Recall that by the weak mixing of the flow, all non-zero time automorphisms are weakly mixing, hence ergodic and therefore disjoint from the identity. Thus, since T_w is disjoint from Id for $w \in P = \{p, -p\}$ and ρ is not product measure, there exist $B_w \in \mathcal{B}, C_w \in \mathcal{C}$ such that

$$(3.4) \quad |\rho(T_{-w}(B_w) \times C_w) - \rho(B_w \times C_w)| > \eta$$

for some $0 < \eta < 1$. There exists $0 < \epsilon < \frac{c\eta}{1000}$ such that

$$\max(|\mu(V_\epsilon^1(B_w)) - \mu(B_w)|, |\nu(V_\epsilon^2(C_w)) - \nu(C_w)|) < \eta/32.$$

Since ρ is a joining, by the triangle inequality, for each $t \in \mathbb{R}$, we have

$$(3.5) \quad |\rho(T_{-t}V_\epsilon^1(B_w) \times V_\epsilon^2(C_w)) - \rho(T_{-t}B_w \times C_w)| < \frac{\eta}{16}.$$

By applying the pointwise ergodic theorem to the joining flow $(T_t \times S_t, \rho)$ and the sets $T_{-w}B_w \times C_w$ and $T_{-w}V_\epsilon^1(B_w) \times V_\epsilon^2(C_w)$ for $w \in P$, there exist $N_0 \in \mathbb{N}$, $\kappa > 0$ (which we can assume additionally to be of the form $\kappa = \kappa(\epsilon)$ as in the assumptions of our theorem) and a set $U_1 \in \mathcal{B} \otimes \mathcal{C}$ with $\rho(U_1) > (1 - \frac{c}{100})\rho(X \times Y)$ (recall that c comes from our assumption) such that for every $L, M \geq N_0$ with $\frac{L}{M} \geq \kappa$ and $w \in P \cup \{0\}$, we have

$$(3.6) \quad \left| \frac{1}{L} \int_M^{M+L} \mathbb{1}_{T_{-w}B_w \times C_w}(T_tx, S_ty) dt - \rho(T_{-w}B_w \times C_w) \right| < \frac{\eta}{16},$$

$$(3.7) \quad \left| \frac{1}{L} \int_M^{M+L} \mathbb{1}_{T_{-w}V_\epsilon^1(B_w) \times V_\epsilon^2(C_w)}(T_tx, S_ty) dt - \rho(T_{-w}V_\epsilon^1(B_w) \times V_\epsilon^2(C_w)) \right| < \frac{\eta}{16}$$

and

$$(3.8) \quad \left| \frac{1}{L} \int_M^{M+L} \mathbb{1}_{T_{-w}B_w \times C_w}(T_{-t}x, S_{-t}y) dt - \rho(T_{-w}B_w \times C_w) \right| < \frac{\eta}{16},$$

$$(3.9) \quad \left| \frac{1}{L} \int_M^{M+L} \mathbb{1}_{T_{-w}V_\epsilon^1(B_w) \times V_\epsilon^2(C_w)}(T_{-t}x, S_{-t}y) dt - \rho(T_{-w}V_\epsilon^1(B_w) \times V_\epsilon^2(C_w)) \right| < \frac{\eta}{16}$$

whenever $(x, y) \in U_1$. Let $U_2 := U_1 \cap (X \times Z)$, where $Z = Z(\epsilon, N_0)$ comes from our assumptions. Then $\rho(U_2) > (1 - c/50)\rho(X \times Y)$. Note also that since $X \times Y$ is σ -compact, measure ρ is regular and hence, we can additionally assume that U_2 is compact. Define $\text{proj} : X \times Y \rightarrow X$, $\text{proj}(x, y) = x$. Then the fibers of proj are σ -compact, and since U_2 is compact, the fibers of the map $\text{proj}|_{U_2} : U_2 \rightarrow \text{proj}(U_2) \subset X$ are also σ -compact and $\text{proj}(U_2)$ is also compact. Thus, by Kunugui's selection theorem (see e.g. [18], Thm. 4.1), it follows that there exists a measurable (selection) $y : \text{proj}(U_2) \rightarrow X \times Y$ such that $(x, y(x)) \in U_2$. Note that $\mu(\text{proj}(U_2)) \geq \rho(U_2) > (1 - c/50)\mu(X)$. By Luzin's theorem there exists $X_{\text{cont}} \subset \text{proj}(U_2)$, $\mu(X_{\text{cont}}) \geq (1 - c/50)\mu(X)$ such that y is uniformly continuous on X_{cont} . Finally, set

$$\tilde{U} := U_2 \cap (X_{\text{cont}} \times Y).$$

We have $\rho(\tilde{U}) > (1 - c/10)\rho(X \times Y)$. Moreover, if $U_X := \text{proj}(\tilde{U})$ then also $\mu(U_X) \geq \rho(\tilde{U}) > (1 - c/10)\rho(X \times Y)$. Hence, by the definitions of sequences (A_k) and $(E_k) = (E_k(\epsilon))$, it follows that there exists $k_0 = k_0(\epsilon)$ such that for $k \geq k_0$, we have

$$(3.10) \quad \mu(A_{-k}(U_X \cap X_k) \cap (U_X \cap X_k) \cap E_k) > 0.$$

Let $\delta = \delta(\epsilon, N_0)$ come from the assumptions of our theorem. By the uniform continuity of $y : X_{\text{cont}} \rightarrow Y$ it follows that there exists $0 < \delta' < \delta$ such that $d_1(x_1, x_2) < \delta'$ implies $d_2(y(x_1), y(x_2)) < \delta$ for each $x_1, x_2 \in X_{\text{cont}}$. Since $A_k \rightarrow Id$ uniformly and $\tilde{U} \subset X_{\text{cont}} \times Y$, there exists $k_1 = k_1(\epsilon)$ such that for $k \geq k_1$, $d_2(y(x), y(A_k x)) < \delta$ for $x \in X_k \cap X_{\text{cont}}$. Fix $k \geq \max(k_0, k_1 + 1)$ (so that $d_1(A_k, Id) < \delta'$). Let $x \in A_{-k}(U_X \cap X_k) \cap (U_X \cap X_k) \cap E_k$. Such a point does exist in view of (3.10). Set $x' = A_k x$, $y = y(x)$, $y' = y(x')$. By definition, $(x, y), (x', y') \in \tilde{U}$ and $d_2(y, y') < \delta$ and all other assumptions of our theorem are satisfied for $(x, y), (x', y')$ (so that we obtain M, L, p depending on (x, y) and (x', y') satisfying (3.2) or (3.3)).

We will assume that (3.2) holds and will get a contradiction using (3.6) and (3.7). If (3.3) holds, we argue analogously using (3.8) and (3.9). We claim that

$$(3.11) \quad \rho(T_{-w}(B_w) \times C_w) > \rho(B_w \times C_w) - \frac{\eta}{2}.$$

Indeed, in view of (3.5), (3.11) follows by showing that

$$\rho(T_{-w}(V_\epsilon^1(B_w)) \times V_\epsilon^2(C_w)) > \rho(B_w \times C_w) - \frac{\eta}{4}.$$

Using (3.2) (for $w = p$), (3.6) (for $w = 0$) and $(x, y) \in \tilde{U} \subset U_1$ (and remembering that for U in (3.2), we have $|U| > (1 - \epsilon)L$ as (a, U, c) is ϵ -good), we obtain that

$$(3.12) \quad \frac{1}{L} \int_M^{M+L} \mathbb{1}_{V_\epsilon^1(B_w) \times V_\epsilon^2(C_w)}(T_{a(t)+w}x', S_{a(t)}y') dt > \frac{1}{L} \int_M^{M+L} \mathbb{1}_{B_w \times C_w}(T_t x, S_t y) dt > \rho(B_w \times C_w) - \epsilon - \frac{\eta}{16}.$$

Hence to complete the proof of claim (3.11), it is enough to show that

$$(3.13) \quad \frac{1}{L} \int_M^{M+L} \mathbb{1}_{V_\epsilon^1(B_w) \times V_\epsilon^2(C_w)}(T_{a(t)+w}x, S_{a(t)}y) dt < \rho(T_{-w}V_\epsilon^1(B_w) \times V_\epsilon^2(C_w)) + \frac{\eta}{8}.$$

Since (a, U, c) is ϵ -good, we get by Remark 3.1 (for $f(t) = \mathbb{1}_{V_\epsilon^1(B_w) \times V_\epsilon^2(C_w)}(T_{t+w}x, S_t y)$)

$$\begin{aligned} \frac{1}{L} \int_M^{M+L} \mathbb{1}_{V_\epsilon^1(B_w) \times V_\epsilon^2(C_w)}(T_{a(t)+w}x, S_{a(t)}y) dt &\leq \\ \frac{1}{L} \int_M^{M+L} \mathbb{1}_{V_\epsilon^1(B_w) \times V_\epsilon^2(C_w)}(T_{s+w}x, S_s y) ds + 9\epsilon/c &\leq \\ \rho(T_{-w}V_\epsilon^1(B_w) \times V_\epsilon^2(C_w)) + \eta/16 + 9\epsilon/c, \end{aligned}$$

where the latter inequality follows from (3.7) for $w \in P$ remembering that $(x, y) \in \tilde{U} \subset U_1$. This completes the proof of (3.13) and hence also of (3.11) since $9\epsilon/c < \eta/16$.

Now, reasoning in a similar way, we get

$$(3.14) \quad \rho(T_{-w}(B_w) \times C_w) < \rho(B_w \times C_w) + \frac{\eta}{2},$$

so putting together (3.11) and (3.14) yields $|\rho(T_{-w}(B_w) \times C_w) - \rho(B_w \times C_w)| < \frac{\eta}{2}$. This however contradicts (3.4). \square

4. DISJOINTNESS FOR TIME CHANGES OF HOROCYCLE FLOWS (PROOF OF THEOREM 1.1)

In this section, as a first application of the disjointness criterion given by Theorem 3.1, we show how it can be used to give a proof of Theorem 1.1 on disjointness of rescalings of smooth time changes of horocycle flows. We first state two results which provide the main ingredients for the proof, namely Proposition 4.1 in Section 4.1, which follows from the form of shearing (in the Ratner property) of time-changes of horocycle flow (and whose proof is postponed to Appendix A), and Lemma 4.2 in Section 4.2, which follows from the work of Bufetov and Forni [6] (see Appendix B). Theorem 1.1 is then proved in Section 4.3, exploiting these two as ingredients and the disjointness criterion.

4.1. Preliminaries and shearing properties of time changes of horocycle flows.

The Lie algebra $sl(2, \mathbb{R})$ is generated by U, V, X , where

$$U := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad V := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad X := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

and X generates the geodesic flow (g_t) , U is the generator of the horocycle flow (h_t) and V generates the opposite horocycle flow (u_t) .

For two points $x, y \in M$ which are sufficiently close, $d_M(x, y) \leq \epsilon_0$ (for some ϵ_0 depending on M), let d_X , d_V and d_U denote distances along, respectively, the geodesic, the opposite horocycle and the horocycle (the distances are measured locally in the Lie algebra). For a function $\xi \in C^1(M)$ and an element $L \in sl(2, \mathbb{R})$, we denote by $L\xi$ the derivative of ξ in direction L .

In this section $\tau \in C^1(M)$ is fixed and we consider (\tilde{h}_t^τ) . For simplicity, we will drop τ from the notation and denote the time change simply by (\tilde{h}_t) . Recall that $u = u(t, x)$ is given by (2.4) (for $v = \tau$).

Take $x, y \in M$ with $d_M(x, y) < \epsilon_0$. Using local coordinates, it follows that

$$x = \exp(d_U(x, y)U) \exp(d_X(x, y)X) \exp(d_V(x, y)V)y,$$

where $d_W(x, y) < 2\epsilon_0$ for $W \in \{U, X, V\}$. We have the following observation, which is a straightforward consequence of the Taylor formula:

Lemma 4.1. *For every $\epsilon > 0$ there exists $\delta > 0$ such that for every $x, y \in M$, $d_M(x, y) < \delta$, we have*

$$|\tau(x) - \tau(y) - \sum_{W \in \{U, X, V\}} d_W(x, y)(W\tau)(x)| < \epsilon \sum_{W \in \{U, X, V\}} d_W(x, y).$$

We will also use the following matrix presentation: if $d_M(x, y) < \epsilon_0$, then (by the right invariance of d_G) there are (unique) small $s = s(x, y)$, $r = r(x, y)$, $\bar{v} = \bar{v}(x, y)$ satisfying

$$(4.1) \quad xy^{-1} = h_{\bar{v}} \begin{pmatrix} e^s & 0 \\ r & e^{-s} \end{pmatrix}.$$

Remark 4.1. Notice that for every $\epsilon > 0$ there exists $\delta > 0$ such that if $d_M(x, y) < \delta$ then

$$\max \left(\left| \frac{d_U(x, y)}{|\bar{v}|} - 1 \right|, \left| \frac{d_X(x, y)}{|s|} - 1 \right|, \left| \frac{d_V(x, y)}{|r|} - 1 \right| \right) < \epsilon^3.$$

Let $\chi_{x, y} : \mathbb{R} \rightarrow \mathbb{R}$ be given by

$$(4.2) \quad \chi_{x, y}(t) = e^{-2s}t - e^{-3s}rt^2.$$

We have the following crucial Proposition 4.1, which provides an exact formula for the amount of splitting in a time-changed flow. This proposition provides a stronger form of

Ratner property (indeed, it implies the Ratner property, see Remark 4.2) in which points are allowed to diverge by an unbounded amount.

In what follows, we set $0^{-1} := +\infty$. Let $A_x(T)$ be such that

$$(4.3) \quad \chi_{x,y}(u(T, x)) = u(\chi_{x,y}(T) + A_x(T), y).$$

(In fact, $A_x(T)$ depends also on y , so $A_x(T) = A_{x,y}(T)$ and this number, or rather $\chi_{x,y}(T) + A_x(T)$, is uniquely determined by $\chi_{x,y}(u(T, x))$ and y , cf. (2.4).)

Proposition 4.1. *Fix $K \geq 1$. For every $\epsilon \in (0, K^{-3})$, there exist $N_\epsilon > 0$ and $\bar{\delta} = \bar{\delta}(\epsilon)$ such that for every x, y satisfying $\max(|r|, |s|, |\bar{v}|) < \bar{\delta}$ (cf. (4.1)) and every $T \in \mathbb{R}$ with $|T| \in [N_\epsilon, K|r|^{-1/2}]$, for $A_x(T)$ defined as above, we have*

$$(4.4) \quad \left| A_x(T) + e^{-2s} \int_0^{u(T,x)} (\tau - \tau \circ g_{-s})(h_t x) dt \right| \leq \epsilon.$$

Moreover, we have

$$(4.5) \quad d_M(\tilde{h}_T x, \tilde{h}_{\chi_{x,y}(T) + A_x(T)} y) \leq \epsilon.$$

Remark 4.2. Proposition 4.1, in particular, implies the Ratner property for the time-changed flow (\tilde{h}_t) (see S 2.5 for the definition). In order to see this, let $\tilde{T} := \min(|r|^{-1/2}, |s|^{-1})$. Notice that by Lemma 4.1, for every $T = O(\tilde{T})$, we have

$$\begin{aligned} \left| \int_0^{u(T,x)} (\tau - \tau \circ g_s)(h_t x) dt \right| &= \left| \int_0^{u(T,x)} s(X\tau)(h_t x) + O(\epsilon^3 |s|) dt \right| \leq \\ &|s| \left| \int_0^{u(T,x)} (X\tau)(h_t x) dt \right| + O(\epsilon^3) = O(\epsilon^3). \end{aligned}$$

Therefore, by (4.4), we have $|A_x(T)| = O(\epsilon^2)$ (we use Proposition 4.1 with ϵ replaced by ϵ^2) and therefore, by (4.5),

$$d_M(\tilde{h}_T x, \tilde{h}_{\chi_{x,y}(T)} y) = O(\epsilon^2),$$

for every $T = O(\tilde{T})$. Recall that $\chi_{x,y}(t) = e^{-2st} - e^{-3s} r t^2$ and hence there exists $T_0 = O(\tilde{T})$ such that $\chi(T_0) = T_0 \pm 1$ and for every $t \in [T_0, (1 + \epsilon^4)T_0]$, $|\chi(t) - \chi(T_0)| = O(\epsilon^2)$. Hence, for $t \in [T_0, (1 + \epsilon^4)T_0]$, we have

$$d_M(\tilde{h}_t x, \tilde{h}_{t \pm 1} y) < \epsilon,$$

what finishes the proof of Ratner property.

The proof of Proposition 4.1, which is a little long and tedious, is postponed to Appendix A.

4.2. Deviations of ergodic averages estimates. In order to prove Theorem 1.1, in addition to Proposition 4.1 stated in the previous section, we will also need the following estimates on ergodic averages for time-changes of a horocycle flow which can be deduced from the work of Bufetov and Forni [6] (see Appendix B).

Recall that from the representation theory of $SL(2, \mathbb{R})$ it follows that the space $L_0^2(M)$ of zero mean square-integrable functions has a decomposition into irreducible (for the regular representation of $SL(2, \mathbb{R})$) components, parametrized by the eigenvalues of the Casimir operator \square (listed with multiplicities) of the following form:

$$L_0^2(M) = \bigoplus_{\mu \in \text{Spec}(\square) \setminus \{0\}} H_\mu = \mathcal{H}_p \oplus \mathcal{H}_c \oplus \mathcal{H}_d,$$

where

$$\mathcal{H}_p = \bigoplus_{\substack{\mu \in \text{Spec}(\square), \\ \mu \geq 1/4}} H_\mu, \quad \mathcal{H}_c = \bigoplus_{\substack{\mu \in \text{Spec}(\square), \\ \mu \in (0, 1/4)}} H_\mu, \quad \mathcal{H}_d = \bigoplus_{\substack{\mu \in \text{Spec}(\square), \\ \mu = -n^2 + n, n \in \mathbb{Z} \setminus \{0\}}} H_\mu.$$

It follows from [12] that the cocycles generated by the functions (cf. 2.3 with R_s replaced by h_s) supported on $\mathcal{H}_d \setminus H_0$ are coboundaries for (h_t) . We will consider time-changes which belong to the class $B^+(M)$ defined as follows. Let $W^\alpha(M) = W^{\alpha, \alpha}(M) \subset L^2(M)$ denote the standard Sobolev space. Let $B(M)$ be given by

$$B(M) = W^6(M) \setminus \mathcal{H}_d \subset L_0^2(M).$$

Thus, the functions in $B(M)$ are not fully contained in the discrete series, or equivalently, in virtue of the above decomposition of $L_0^2(M)$, the elements of $B(M)$ are those functions in $W^6(M)$ projecting non-trivially on $\mathcal{H}_c \oplus \mathcal{H}_d$. Finally, we let $B^+(M)$ consist of those *positive* functions τ (which can then be taken as roof functions) which are of the form $\tau = c + \tau'$, where $\tau' \in B(M)$ and $c \in \mathbb{R}$. We call functions with non-trivial support on \mathcal{H}_c of *type I* and functions supported on \mathcal{H}_p of *type II*.

The following lemma is a consequence of the work by Bufetov and Forni [6], more precisely, of Lemma B.1, which is deduced from [6] in Appendix B (for β_τ , $\beta_\tau^{(1/4)}$ and \mathbf{v} , see Lemma B.1).

Lemma 4.2 (Consequence of Bufetov-Forni [6]). *For every $\epsilon > 0$, there exists $T_\epsilon > 0$ such that*

(T1) *if τ is of type I then there exist $\alpha_\tau \in (0, 1)$, $c_\tau \in \mathbb{R} \setminus \{0\}$ such that for $|s| < T_\epsilon^{-1}$ and $|T| = O(|s|^{-1/\alpha_\tau})$, we have*

$$\left| \int_0^T (\tau - \tau \circ g_s)(h_t x) dt - s c_\tau T^{\alpha_\tau} \beta_\tau(g_{\log T} x) \right| < \epsilon^2;$$

(T2) *if τ is of type II and $\beta_\tau^{(1/4)}$ vanishes identically then for $|s| < T_\epsilon^{-1}$ and $|T| = O(|s|^{-2})$, we have*

$$\left| \int_0^T (\tau - \tau \circ g_s)(h_t x) dt - s T^{1/2} \left(\frac{d}{ds} \beta_\tau \right) (0, \mathbf{v} \log T, g_{\log T} x) \right| < \epsilon^2;$$

(T3) *if τ is of type II and $\beta_\tau^{(1/4)}$ does not vanish identically, then for $|s| < T_\epsilon^{-1}$ and $|T| = O(\frac{|s|^{-2}}{\log^2 s})$, we have*

$$\left| \int_0^T (\tau - \tau \circ g_s)(h_t x) dt + \frac{s}{2} T^{1/2} \log T \beta_\tau^{(1/4)}(g_{\log T} x) \right| < \epsilon^2.$$

Proof. The proof in case (T1) is a straightforward consequence of the first part of Lemma B.1, since $\phi_\tau(s) = \phi'_\tau(0)s + O(s^2)$ and hence (setting $c_\tau := \phi'_\tau(0)$)

$$\phi_\tau(s) T^{\alpha_\tau} \beta_\tau(g_{\log T} x) = c_\tau s T^{\alpha_\tau} \beta_\tau(g_{\log T} x) + O(s^2 T^{\alpha_\tau}),$$

where $O(s^2 T^{\alpha_\tau}) = O(|s|)$. For (T2), we have

$$\beta_\tau(s, \mathbf{v} \log T, g_{\log T} x) = s \left(\frac{d}{ds} \beta_\tau \right) (0, \mathbf{v} \log T, g_{\log T} x) + O(s^2)$$

and a reasoning analogous to the above one applies. An analogous reasoning (expanding $e^{-s/2} - 1$ at 0) gives (T3). This finishes the proof. \square

In the next section, we will also make use of the following observation.

Remark 4.3. Recall that the linear space $\mathcal{P}([0, 1]) := \{p : [0, 1] \rightarrow \mathbb{R} : p \text{ is a quadratic polynomial}\}$ is finite dimensional and hence any two norms on this space are equivalent. In particular, it follows that there exists a constant $C_{\mathcal{P}}$ such that for every $U > 0$ and every quadratic polynomial $w : \mathbb{R} \rightarrow \mathbb{R}$, $w = at^2 + bt + c$,

$$C_{\mathcal{P}}^{-1} \max(|a|U^2, |b|U, |c|) < \sup_{t \in [0, U]} |w(t)| < C_{\mathcal{P}} \max(|a|U^2, |b|U, |c|).$$

4.3. Proof of Theorem 1.1. In this section we prove Theorem 1.1, by showing how the assumptions of the disjointness criterion can be verified using Proposition 4.1 and Lemma 4.2.

Proof of Theorem 1.1. We will first give a proof when the assumptions of (T2) are satisfied, in fact, this is the most complicated case. We will then state what changes are needed if τ satisfies (T1) or (T3). Fix p, q and assume WLOG that $0 < p < q$. Let $A := \{\mathbf{v}t : t \in \mathbb{R}\} \subset \mathbb{T}^\infty$. Notice that if $c' > 0$ is small enough, then for some $d_{p,q} > 0$, we have

$$(4.6) \quad \|d/ds(\beta_\tau)(0, \cdot, \cdot)\|_{C^0(A \times M)} \left((p^{-1/2} - q^{-1/2}) - \frac{c'}{p^{1/2}} \right) > 2d_{p,q}.$$

Set

$$(4.7) \quad L := \{(a, x) \in A \times M : |d/ds(\beta_\tau)(0, a, x)| > \|d/ds(\beta_\tau)(0, \cdot, \cdot)\|_{C^0(A \times M)} - c'\}.$$

By the continuity of $d/ds(\beta_\tau)(0, \cdot, \cdot)$ there exists a set $S \times R \subset L$, with $\lambda_\infty(S)\mu(R) =: c > 0$, where λ_∞ stands for Haar measure on \mathbb{T}^∞ . Moreover, since L is open, we can choose S being open.

We will show that the assumptions of Theorem 3.1 are satisfied. Let $P = \{-r_{p,q}, r_{p,q}\}$, where $r_{p,q} > 0$ is a small constant to be specified later. Let $(n_k) \subset \mathbb{R}$ be an increasing sequence going to ∞ such that $\mathbf{v} \log(pn_k^2) \in S$ (such (n_k) exists since S is open and the orbit of 0 under the linear flow in direction \mathbf{v} is dense in A). Define $X_k := M$ and $A_k(x) = g_{1/n_k}(x)$. Then obviously $A_k \rightarrow \text{Id}$ uniformly. Fix $\epsilon > 0$ and $N \in \mathbb{N}$. Let $E_k = g_{-\log(pn_k^2)}(R)$; then $\mu(E_k) = \mu(R) > c$.

Let $\kappa = \kappa(\epsilon) = \epsilon^4$, $Z(\epsilon, N) = M$ and $\delta = \delta(\epsilon, N) = \min(\epsilon^{10}, T_\epsilon^{-10}, \bar{\delta})$, where $\bar{\delta} = \bar{\delta}(\epsilon)$ comes from Proposition 4.1 and T_ϵ comes from Lemma 4.2. Take k such that $d(A_k, \text{Id}) < \delta$, $x \in E_k$, $x' = A_k x$ and $y, y' \in M$ so that $\max(d_M(x, x'), d_M(y, y')) < \delta$. By the definition of A_k , it follows that $r(x, x') = 0$, $\bar{v}(x, x') = 0$ and $s(x, x') = 1/n_k$ (see (4.1)).

Set

$$(4.8) \quad \bar{T} := \min(n_k^2, s(y, y')^{-2}, |r(y, y')|^{-1/2}).$$

We claim now that

$$(4.9) \quad \left| \frac{1}{p} \int_0^{u(pn_k^2, x)} (\tau - \tau \circ g_{n_k^{-1}})(h_t x) dt - \frac{1}{q} \int_0^{u(qn_k^2, y)} (\tau - \tau \circ g_{n_k^{-1}})(h_t y) dt \right| \geq d_{p,q}.$$

Indeed, notice first that by (T2), for every $W = O(n_k^2)$ and every $z \in M$, we have (since $|u(T, \cdot) - T| < \epsilon^4 T$)

$$\begin{aligned} & \left| \int_0^{u(W, z)} (\tau - \tau \circ g_{n_k^{-1}})(h_t z) dt - \int_0^W (\tau - \tau \circ g_{n_k^{-1}})(h_t z) dt \right| \leq \\ & \left| \int_0^{u(W, z) - W} (\tau - \tau \circ g_{n_k^{-1}})(h_t(h_W z)) dt \right| = O(n_k^{-1}(\epsilon^4 n_k^2)^{1/2}) = O(\epsilon^2). \end{aligned}$$

Hence, (4.9) is by (T2), up to $O(\epsilon^2)$, equal to

$$\left| \frac{1}{p} n_k^{-1} (pn_k^2)^{1/2} (d/ds\beta_\tau)(0, \mathbf{v} \log(pn_k^2), g_{\log(pn_k^2)}x) - \frac{1}{q} n_k^{-1} (qn_k^2)^{1/2} (d/ds\beta_\tau)(0, \mathbf{v} \log(qn_k^2), g_{\log(qn_k^2)}y) \right|.$$

Since $x \in E_k$, we have $(\mathbf{v} \log(pn_k^2), g_{\log(pn_k^2)}x) \in S \times R \subset L$ and therefore, by (4.7) and the triangle inequality, the above expression is larger than

$$\|d/ds(\beta_\tau)(0, \cdot, \cdot)\|_{C^0(A \times M)} \left(|p^{-1/2} - c_{p,q}| - q^{-1/2} \right),$$

which, by (4.6), gives (4.9).

By (4.5), in Proposition 4.1 (with K to be specified at the end of the proof), for x, x' and then for y, y' , using (4.8), for $N_\epsilon \leq t \leq K\bar{T}$, we get

$$(4.10) \quad d_M(\tilde{h}_{pt}x, \tilde{h}_{\chi_{x,x'}(pt)+A_x(pt)}x') \leq \epsilon^2 \quad \text{and} \quad d_M(\tilde{h}_{qt}y, \tilde{h}_{\chi_{y,y'}(qt)+A_y(qt)}y') \leq \epsilon^2.$$

Define $a(t) := \frac{1}{q}(\chi_{y,y'}(qt) + A_y(qt))$. By the definition of $\chi(\cdot)$ and $A_y(\cdot)$ (from (4.3), we have that $A_y(\cdot)$ is smooth on every interval $I \subset [0, K\bar{T}]$, it follows that a satisfies (SAL) from Definition 3.1 for every interval $I = [\eta_1\bar{T}, \eta_2\bar{T}] \subset [0, K\bar{T}]$.³

From (4.10), for all $N_\epsilon \leq t \leq \bar{T}$, we have

$$d_M(\tilde{h}_{qt}y, \tilde{h}_{qa(t)}y') \leq \epsilon^2.$$

Let moreover $b(t) = \frac{1}{p}(\chi_{x,x'}(pt) + A_x(pt))$. Then, analogously, by (4.10), we have

$$d_M(\tilde{h}_{pt}x, \tilde{h}_{pb(t)}x') \leq \epsilon^2.$$

So to finish the proof of (3.2), it is enough to show that there exists $t_0 \in [N_\epsilon, K\bar{T}]$ such that

$$(4.11) \quad |a(t_0) - b(t_0) - r_{p,q}| < \epsilon^2,$$

for some $r_{p,q} \neq 0$ and, for every $t \in [0, \kappa t_0]$, we have

$$(4.12) \quad |a(t+t_0) - b(t+t_0) - a(t_0) - b(t_0)| < \epsilon^2.$$

By the definition of $\chi_{x,x'}(t)$ and $\chi_{y,y'}(t)$, we get

$$(4.13) \quad a(t) - b(t) = pol(t) + sub(t),$$

where

$$pol(t) = (e^{-2s(y,y')} - e^{-2s(x,x')})t - qe^{-3s(y,y')}r(y,y')t^2, \quad \text{and} \quad sub(t) = \frac{A_x(pt)}{p} - \frac{A_y(qt)}{q}$$

(here we use the fact that $r(x, x') = 0$). Notice that by (4.4) for x and by (T2), for every $t_0 \in [0, K\bar{T}]$ and every $t \in [0, \kappa t_0]$, we have (recalling that $s(x, x') = n_k^{-1}$)

$$\frac{1}{p}|A_x(p(t+t_0)) - A_x(pt_0)| = e^{-2s} \int_0^{u(p(t+t_0), x) - u(pt_0, x)} (\tau - \tau \circ g_{s(x, x')})(h_t(h_{u(pt_0, x)}x)) d\tau + O(\epsilon^2) \leq$$

$$s(x, x') (u(p(t+t_0), x) - u(pt_0, x))^{1/2} \|d/ds(\beta_\tau)(0, \cdot, \cdot)\|_{C^0(A \times M)} + O(\epsilon^2) = O(\kappa) + O(\epsilon^2),$$

³Indeed, in view of (4.3),

$$a(t) = \frac{1}{q} \int_0^{\chi_y(u(qt, y))} \tau(h_\theta y) d\theta = \frac{1}{q} \int_0^{\chi_y(u(qt, y)) - u(qt, y)} \tau(h_\theta y) d\theta + t.$$

Now, $\chi_y(u(qt, y)) - u(qt, y) = O(|s|t) = O(|s|\bar{T}) \leq \epsilon\bar{T}$ and since the integrand is bounded, we only need to show that $\frac{d}{dt}(\chi_y(u(qt, y)) - u(qt, y)) = O(\epsilon^2)$. The latter follows from the fact that $u'(qt, y)$ is bounded and $\chi_y'(u(qt, y)) - 1 = e^{-2s} - 1 + O(r\bar{T})$.

where the (only) inequality above holds since $|u(p(t+t_0), x) - u(pt_0, x)| = O(t) = \kappa O(\bar{T}) = \kappa O(s(x, x')^2)$. Therefore, $\frac{1}{p}|A_x(p(t+t_0)) - A_x(pt_0)| = O(\kappa) + O(\epsilon^2)$. An analogous reasoning for y shows that $\frac{1}{q}|A_y(q(t+t_0)) - A_y(qt_0)| = O(\kappa) + O(\epsilon^2)$. Hence, for every $t_0 \in [0, K\bar{T}]$ and every $t \in [0, \kappa t_0]$, we have

$$(4.14) \quad |sub(t+t_0) - sub(t_0)| = O(\epsilon^2).$$

By definition, the functions $a(\cdot)$ and $b(\cdot)$ are continuous on $[0, K\bar{T}]$ and therefore to prove (4.11), it is enough to show that there exists $t_1 < K\bar{T}$ such that

$$(4.15) \quad |a(t_1) - b(t_1)| \geq 2|r_{p,q}|.$$

We consider two cases:

A. There exists $t' \in [0, 2\bar{T}]$ such that $|pol(t')| > d_{p,q}/2$ (see (4.9)). Let then $T' > 0$ be the smallest number for which $|pol(T')| = d_{p,q}/2$. By Remark 4.3, for the polynomial $k \mapsto pol(kT')$ and for every $\tilde{K} > 2$, we have

$$(4.16) \quad \sup_{k \in [0, \tilde{K}]} |pol(kT')| \geq C_{\mathcal{P}}^{-2} \tilde{K} \frac{d_{p,q}}{2}.$$

Moreover, by (4.4) for x , by (T2) and (4.8) (since $T' < 2\bar{T}$), for $k \in [0, K]$, we have

$$(4.17) \quad \frac{1}{p}|A_x(pkT')| = O(\epsilon^2) + p^{-1}|s(x, x')|(pkT' + O(\epsilon^4 pkT'))^{1/2} \|(d/d\theta \beta_\tau)(0, \cdot, \cdot)\|_{C^0(A \times M)} \leq \frac{2k^{1/2}}{p^{-1/2}} \|(d/d\theta \beta_\tau)(0, \cdot, \cdot)\|_{C^0(A \times M)},$$

and similarly

$$(4.18) \quad \frac{1}{q}|A_y(qkT')| \leq \frac{2k^{1/2}}{q^{-1/2}} \|(d/d\theta \beta_\tau)(0, \cdot, \cdot)\|_{C^0(A \times M)}.$$

Notice that if $\tilde{K} > 0$ is such that

$$C_{\mathcal{P}}^{-2} \tilde{K} \frac{d_{p,q}}{4} > 2\tilde{K}^{1/2} \left(\frac{1}{p^{-1/2}} + \frac{1}{q^{-1/2}} \right) \|(d/ds \beta_\tau)(0, \cdot, \cdot)\|_{C^0(A \times M)},$$

then by (4.16), (4.17) and (4.18) it follows that there exists $k \in [0, \tilde{K}]$ such that (see (4.13))

$$|a(kT') - b(kT')| \geq |pol(kT')| - |sub(kT')| \geq C_{\mathcal{P}}^{-2} \tilde{K} \frac{d_{p,q}}{4}$$

what finishes the proof of (4.15) (and hence also the proof of (4.11), where $t_0 \leq kT'$ if $|r_{p,q}|$ is small enough and $\tilde{K} < K$. Moreover, by Remark 4.3 and the definition of T' , it follows that for every $t \in [0, \kappa t_0]$, we have

$$|pol(t+t_0) - pol(t_0)| = O(\epsilon^2).$$

This and (4.14) (see (4.13)) finishes the proof of (4.12) and the proof of Theorem 1.1 is complete in this case.

B. For every $t' \in [0, 2\bar{T}]$, we have $|pol(t')| < \frac{d_{p,q}}{2}$. This, by Remark 4.3, in particular means that

$$(4.19) \quad |e^{-2n_k^{-1}} - e^{-2s(y, y')}| \leq \frac{C_{\mathcal{P}} d_{p,q}}{2\bar{T}}$$

and

$$(4.20) \quad |r(y, y')| \leq \frac{C_{\mathcal{P}} d_{p,q}}{4q\bar{T}^2}.$$

We claim moreover that

$$(4.21) \quad 2\bar{T} \geq n_k^2.$$

Indeed, if not then by (4.8), $\bar{T} = \min(|r(y, y')|^{-1/2}, s(y, y')^{-2})$. If $\bar{T} = s(y, y')^{-2}$, then by (4.21), we get $-2n_k^{-1} < -2^{1/2}|s(y, y')|$. But then (by Taylor's formula)

$$|e^{-2n_k^{-1}} - e^{-2s(y, y')}| \geq e^{-2s(y, y')} - e^{-2^{1/2}s(y, y')} \geq \frac{2 - 2^{1/2}}{2}|s(y, y')|$$

which is a contradiction with (4.19). On the other hand, if $\bar{T} = |r|^{-1/2}$, then we get a contradiction with (4.20) (by making $d_{p,q}$ smaller if necessary). Hence (4.21) holds. Notice that by (4.19) and (4.20) (see also Remark 4.3), for every $t_0 \in [0, 2\bar{T}]$ and every $t \in [0, \kappa t_0]$, we have

$$pol(t + t_0) - pol(t_0) = O(\epsilon^2).$$

This and (4.14) (see (4.13)) show that (4.12) holds for every $t_0 \leq 2\bar{T}$. So we only have to show that (4.15) holds for some $t_0 \leq 2\bar{T}$.

Notice that by Lemma 4.1 for $g_{n_k^{-1}}h_t y$ (recall that $n_k^{-1} = s(x, x')$) and $g_{s(y, y')^{-1}}h_t y$, we have

$$\begin{aligned} & \frac{1}{q} \left| \int_0^{u(qn_k^2, y)} (\tau - \tau \circ g_{n_k^{-1}})(h_t y) dt - \int_0^{u(qn_k^2, y)} (\tau - \tau \circ g_{s(y, y')^{-1}})(h_t y) dt \right| = \\ & \frac{1}{q} \left| \int_0^{u(qn_k^2, y)} (s(x, x') - s(y, y'))(X\tau)(g_{s(y, y')^{-1}}h_t y) + O(\epsilon^3|s(x, x') - s(y, y')|) dt \right|. \end{aligned}$$

Moreover, using the equality $g_{s(y, y')^{-1}}h_t y = h_{e^{-2s(y, y')}t} g_{s(y, y')}$, substituting $t' = e^{-2s(y, y')}t$ and using that $\int_M X\tau d\mu = 0$, we get that the above expression is equal to

$$|s(x, x') - s(y, y')| O(\epsilon^3 u(qn_k^2, y)) = |s(x, x') - s(y, y')| O(\epsilon^3 n_k^2) = O(\epsilon^3),$$

where the last equality follows by (4.19) and (4.21). Therefore, by (4.9) and (4.4), it follows that if we set $t_1 = n_k^2$ (we also use (4.21) to be able to use (4.4)), then

$$|sub(t_1)| \geq \frac{2d_{p,q}}{3}.$$

By (4.21), it follows that $t_1 = n_k^2 < 2\bar{T}$ and hence by the assumption in **B.**, we have $|pol(t_1)| < \frac{d_{p,q}}{2}$. This implies that

$$|a(t_1) - b(t_1)| \geq |sub(t_1)| - |pol(t_1)| \geq \frac{d_{p,q}}{6}$$

and this gives (4.15) (if $|r_{p,q}|$ is small enough) and hence also (4.11). This finishes the proof in case τ satisfies (T2).

In case τ satisfies the assumptions of (T1) or (T3), we define analogously

$$L := \{x \in M : |\Xi(x)| > \|\Xi(\cdot)\|_{C^0(M)} - c'\},$$

where $\Xi = \beta_\tau$ in (T1) and $\Xi = \beta_\tau^{(1/4)}$ in (T3), and $\mu(L) > c$.

In case τ satisfies the assumptions of (T1), we set $X_k = M$, $A_k x = g_{k^{-1}}x$ and $E_k = g_{-\log p k^{1/\alpha_\tau}}(L)$; and $X_k = M$, $A_k x = g_{k^{-1}}x$ and $E_k = g_{-\log(pk^2 \log k)}(L)$. We define κ, δ, Z as above. Finally, we set

$$\bar{T} := \min(k^{1/\alpha_\tau}, r^{-1/2}(y, y), s(y, y')^{-1/\alpha_\tau})$$

if τ satisfies (T1) and

$$\bar{T} := \min(k^{1/\alpha_\tau}, r^{-1/2}(y, y), s(y, y')^{-1/\alpha_\tau})$$

if τ satisfies (T3). The rest of the proof follows the lines of the proof in case (T2), i.e. we first show (4.9) with $u(pk^{1/\alpha_\tau})$ and $u(qk^{1/\alpha_\tau})$ in case (T1) and $u(pk^2 \log k)$ and $u(qk^2 \log k)$ in case (T3). The functions $a(\cdot)$ and $b(\cdot)$ are defined by the same expressions. The proof (with these new definitions) follows the same lines as the proof in case (T2). This completes the proof of Theorem 1.1. \square

5. DISJOINTNESS CRITERION FOR SPECIAL FLOWS

In this section we assume that $(T_t) = (T_t^f)$ and $(S_t) = (S_t^g)$ are special flows over ergodic isometries $T \in \text{Aut}(X, \mathcal{B}, \mu, d_1)$, $S \in \text{Aut}(Y, \mathcal{C}, \nu, d_2)$ and under $f \in L_+^1(X, \mathcal{B}, \mu)$, $g \in L_+^1(Y, \mathcal{C}, \nu)$, respectively. Our aim is to give assumptions on the parameters of special flows so that Theorem 3.1 applies.

Then (T_t^f) acts on X^f with metric d_1^f (which is the restriction of the product metric) and (S_t^g) acts on Y^g with metric d_2^g . For $(x, s) \in X^f$ and $t \in \mathbb{R}$, we denote by $n(x, s, t) \in \mathbb{Z}$ the unique number for which

$$f^{(n(x,s,t))}(x) \leq t + s < f^{(n(x,s,t)+1)}(x),$$

i.e.

$$(5.1) \quad T_t^f(x, s) = (T^{n(x,s,t)}x, s + t - f^{n(x,s,t)}(x)).$$

We define $m(y, r, t)$ analogously for $(y, r) \in Y^g$. We assume that

$$(5.2) \quad f \text{ and } g \text{ are bounded away from zero.}$$

Note that

$$(5.3) \quad n(x, s, t + 1) - n(x, s, t) \leq 1 / \min(1, \min f).$$

Proposition 5.1 (Disjointness criterion for special flows). *Let $P = \{-p, p\}$ with $p \neq 0$, $0 < c < 1/2$ and let*

$$\zeta := \frac{\int_X f d\mu}{\int_Y g d\nu}.$$

Assume that we have a sequence of measurable sets

$$(X'_k) \subset X, \quad \mu(X'_k) \rightarrow \mu(X),$$

together with a sequence of automorphisms

$$A'_k \in \text{Aut}(X'_k, \mathcal{B}|_{X'_k}, \mu|_{X'_k}), \quad k \geq 1, \text{ such that } A'_k \rightarrow \text{Id uniformly.}$$

Assume moreover that for every $\epsilon' > 0$ and $N' \in \mathbb{N}$ there exist

$$(E'_k) = (E'_k(\epsilon')) \subset \mathcal{B}, \quad \mu(E'_k) > c \text{ for } k \geq 1,$$

as well as $0 < \kappa' = \kappa'(\epsilon') < \epsilon'$, $\delta' = \delta(\epsilon', N') > 0$ and a set

$$Z' = Z(\epsilon', N') \subset Y, \quad \nu(Z') \geq (1 - \epsilon')\nu(Y),$$

such that for all $y, y' \in Z'$ satisfying $d_2(y, y') < \delta'$, every k such that $d_1(A'_k, \text{Id}) < \delta'$ and every $x \in E'_k \cap X'_k$, $x' := A'_k x$ there are

$$M' \geq N, \quad L' \geq 1, \quad \frac{L'}{M'} \geq \kappa'$$

*and $p \in P$ for which one of the following sets of estimates, **(F)** or **(B)** (where **F** stands for Forward and **B** for Backward), holds:*

(F) Forward control:

(F1) For any t such that $n(x, s, t) \in [M', M' + L']$, we have⁴

$$\max \left(\left| n(x, s, t) \int_X f d\mu - t \right|, \left| m(y, r, t) \int_Y g d\nu - t \right| \right) < t^{1/3};$$

(F2) for $w \in [M', M' + L'] \cap \mathbb{Z}$, we have

$$\left| (f^{(w)}(x) - f^{(w)}(x')) - (g^{([\zeta w])}(y) - g^{([\zeta w])}(y')) - p \right| < \epsilon';$$

(F3) for $w \in [0, \max(1, \zeta)(M' + L')] \cap \mathbb{Z}$, we have

$$|f(T^w x) - f(T^w x')| < \kappa(\epsilon') \text{ and } |g(S^w y) - g(S^w y')| < \kappa(\epsilon');$$

(F4) for every $w, u \in \left[\frac{\zeta M'}{2}, 2\zeta(M' + L') \right]$, $|w - u| \leq M'^{1/2}$, we have

$$\left| (g^{(w)}(y) - g^{(w)}(y')) - (g^{(u)}(y) - g^{(u)}(y')) \right| < \epsilon'.$$

(B) Backward control:

(B1) For any t such that $n(x, s, -t) \in [M', M' + L']$, we have

$$\max \left(\left| n(x, s, -t) - t \int_X f d\mu \right|, \left| m(y, r, -t) - t \int_Y g d\nu \right| \right) < t^{1/3};$$

(B2) for $w \in [M', M' + L'] \cap \mathbb{Z}$, we have

$$\left| (f^{(-w)}(x) - f^{(-w)}(x')) - (g^{([- \zeta w])}(y) - g^{([- \zeta w])}(y')) - p \right| < \epsilon';$$

(B3) for $w \in [0, \max(1, \zeta)(M' + L')] \cap \mathbb{Z}$, we have

$$|f(T^{-w} x) - f(T^{-w} x')| < \kappa(\epsilon') \text{ and } |g(S^{-w} y) - g(S^{-w} y')| < \kappa(\epsilon');$$

(B4) for every $w, u \in \left[\frac{\zeta M'}{2}, 2\zeta(M' + L') \right]$, $|w - u| \leq M'^{1/2}$, we have

$$\left| (g^{(-w)}(y) - g^{(-w)}(y')) - (g^{(-u)}(y) - g^{(-u)}(y')) \right| < \epsilon'.$$

Then (T_t^f) and (S_t^g) are disjoint.

The rest of this section will be devoted to the proof of this criterion, which will be deduced from the disjointness criterion for flows given by Theorem 3.1.

Proof. We will show that the assumptions of Theorem 3.1 (with c replaced by $c/2$) are satisfied. Let

$$C_{f,g} := \max \left(1, \int_X f d\mu, \left(\int_X f d\mu \right)^{-1}, \int_Y g d\nu, \left(\int_Y g d\nu \right)^{-1} \right),$$

$$c_{f,g} := \min(1, \inf_{\mathbb{T}} f, \inf_{\mathbb{T}} g) > 0.$$

Fix $\epsilon > 0$ and $N \in \mathbb{N}$. Set $\epsilon' := \frac{\epsilon^2}{2c_{f,g}}$. Let $\kappa' = \kappa'(\epsilon') > 0$ be obtained from the assumption of our proposition and set

$$(5.4) \quad \kappa := \min \left(\frac{\kappa'}{100C_{f,g}^2}, \frac{1}{100}, \frac{1}{18(c_{f,g}^{-1} + 1) \left(\int f d\mu \right)^{2/3}} \right).$$

By taking ϵ' still smaller, we can assume that

$$(5.5) \quad \frac{4\kappa'^2}{\int f d\mu} < \kappa\epsilon.$$

⁴Here $s, r \in \mathbb{R}$ are any numbers for which $(x, s) \in X^f$ and $(y, r) \in Y^g$.

Let

$$W^f(\epsilon) := \left\{ (x, s) \in X^f : \frac{\epsilon}{100} < s < f(x) - \frac{\epsilon}{100} \right\}$$

and

$$W^g(\epsilon) := \left\{ (y, r) \in Y^g : \frac{\epsilon}{100} < r < g(y) - \frac{\epsilon}{100} \right\}.$$

For $J \subset \mathbb{R}$ let

$$U_J^f(x, s) := \{t \in J : T_t^f(x, s) \in W^f(\epsilon)\}, \quad U_J^g(y, r) := \{t \in J : S_t^g(y, r) \in W^g(\epsilon)\}.$$

By Birkhoff ergodic theorem there exist $T_\epsilon > 0$, $V_1 \subset X^f$, $\mu^f(V_1) \geq (1 - \frac{\epsilon}{2})\mu^f(X^f)$ and $V_2 \subset Y^g$, $\nu^g(V_2) \geq (1 - \frac{\epsilon}{2})\nu^g(Y^g)$ such that for every $(x, s) \in V_1$, $(y, r) \in V_2$, we have

$$(5.6) \quad |U_J^f(x, s)| > \left(1 - \frac{\epsilon}{2}\right) |J| \text{ and } |U_J^g(y, r)| > \left(1 - \frac{\epsilon}{2}\right) |J|,$$

holds for all $J = [U, U + V]$, with $\frac{|V|}{|U|} \geq \kappa'$, $|U| \geq T_\epsilon$.

Define $N' := [\max(2T_\epsilon, \kappa^{-3}N, \epsilon^{-1}, c_{f,g}^{-4})] + 1$. Moreover, by enlarging N' if necessary, we have

$$(5.7) \quad N_1'^{1/2} \leq \frac{\kappa'}{10} \int f d\mu \cdot N_1' \text{ for each } N_1' \geq N'.$$

Let

$$X_k := \{(x, s) \in X^f : x \in X_k', (A_k'x, s) \in X^f\}$$

and set

$$A_k(x, s) := (A_k'x, s)$$

for $(x, s) \in X_k$. Finally, set

$$E_k = E_k(\epsilon) := \left\{ (x, s) \in X^f : x \in E_k'(\epsilon'), (x, s) \in V_1, s < \frac{1}{\epsilon^{j_0}} \right\},$$

where j_0 is chosen in such a way that whenever for a set $C \subset X$, we have $\mu(C) > c$ then $\mu^f(\{(x, s) \in X^f : x \in C, s < 1/\epsilon^{j_0}\}) > \frac{3}{4}c$ (the existence of j_0 follows immediately from the fact that f is in L^1). By the properties of A_k', X_k' and $E_k'(\epsilon')$ it follows that A_k, X_k and E_k satisfy the assumptions of Theorem 3.1. Let

$$Z(\epsilon, N) = \{(y, r) \in Y^g : y \in Z'(\epsilon', N'), r < \epsilon^{-j_0}\} \cap V_2^5$$

and

$$\delta(\epsilon, N) := \min(\delta'(\epsilon', N'), \frac{\epsilon}{1000}).$$

Take k such that $d_1^f(A_k, Id) < \delta$, $(x, s) \in E_k$ and set $(x', s') := A_k(x, s)$, take $(y, r), (y', r') \in Z(\epsilon, N)$ for which $d_2^g((y, r), (y', r')) < \delta$. Then $d_1(A_k', Id) < \delta$, $x \in E_k'$, $x' = A_k'x$, $y, y' \in Z'(\epsilon', N')$ and $d_2(y, y') < \delta'$. Hence A. or B. holds for (x, x') and (y, y') . We will assume that A. holds and show that (3.2) holds for $((x, s), (x', s')), ((y, r), (y', r'))$ (if B. holds we argue analogously to show that (3.3) holds). Since A. holds, we obtain M', L' and p .

Let $M, L \in \mathbb{N}$ be any numbers such that $n(x, s, M) \in [M' + 1, M' + c_{f,g}^{-1} + 1] \cap \mathbb{Z}$, and $n(x, s, M + L) \in [M' + L' - c_{f,g}^{-1} - 1, M' + L' - 1] \cap \mathbb{Z}$. Notice that such M, L always exist. Indeed, this follows by (5.2) and (5.3): for all $(x, s) \in X^f$, $n(x, s, u + 1) - n(x, s, u) \leq c_{f,g}^{-1}$.

We also have:

$$(5.8) \quad \text{if } t \in [M, M + L] \text{ then } n(x, s, t) \in [M', M' + L'].$$

⁵Again, by changing j_0 if necessary, we can assume that $\nu^g(Z(\epsilon, N)) > 1 - \epsilon$.

We will show that $M \geq N$, $\frac{L}{M} \geq \kappa$. By property (F1) (and (5.8)) for $t = M$, it follows that

$$(5.9) \quad 2M \geq M + M^{1/3} \geq n(x, s, M) \int_X f d\mu \geq M' \int_X f d\mu.$$

Hence $M \geq N$ (since $M' \geq N' \geq \kappa^{-3}N$, and use (5.9)). Moreover,

$$(5.10) \quad M' \geq \frac{n(x, s, M)}{2} \geq \frac{M - M^{1/3}}{2 \int_X f d\mu} \geq \frac{M}{3 \int_X f d\mu}^6$$

and by property (F1) of Prop. 5.1 for $t = M + L$, we have

$$(5.11) \quad \frac{M + L}{2} \leq (M + L) - (M + L)^{1/3} \leq n(x, s, M + L) \int_X f d\mu \leq (M' + L') \int_X f d\mu.$$

Again by property (F1) first for $t = M + L$ and then $t = M$, the definition of M and L and κ (to see that $2(c_{f,g}^{-1} + 1) \int f d\mu \leq M^{1/3} < (M + L)^{1/3}$), by (5.11) (and the definition of κ to see that $3(2(M' + L') \int f d\mu)^{1/3} \leq (M' + L')^{1/2}$), (5.7), (5.10), we have

$$\begin{aligned} L &= (M + L) - M \geq \\ & n(x, s, M + L) \int_X f d\mu - (M + L)^{1/3} - n(x, s, M) \int_X f d\mu - M^{1/3} \geq \\ & (n(x, s, M + L) - n(x, s, M)) \int_X f d\mu - 2(M + L)^{1/3} \geq \\ & ((M' + L') - c_{f,g}^{-1} - 1 - M' - c_{f,g}^{-1} - 1) \int_X f d\mu - 2(M + L)^{1/3} \geq L' \int_X f d\mu - 3(M + L)^{1/3} \geq \\ & L' \int_X f d\mu - (M' + L')^{1/2} \geq L' \int f d\mu - \frac{\kappa'}{10} \int f d\mu (L' + M') \geq \\ & \frac{2}{3} L' \int f d\mu - \frac{\kappa'}{10} \int f d\mu \cdot M' \geq \\ & \left(\int f d\mu \cdot \kappa' M' \right) \left(\frac{2}{3} - \frac{1}{10} \right) \geq \frac{1}{2} \int f d\mu \cdot \kappa' M' \geq \\ & \frac{\int f d\mu}{2} \kappa' \frac{M}{3 \int f d\mu} \geq \kappa M, \end{aligned}$$

the last inequality by the definition of κ .

Let $U := U_{[M, M+L]}^f(x, s) \cap U_{[M, M+L]}^g(y, r)$. By (5.6), we have $|U| \geq (1 - \epsilon)L$. By the definition of U (see the definitions of U_J^f, U_J^g and W^f, W^g), it follows that U is a disjoint union of intervals, $U = \bigcup_{i=1}^v (c_i, d_i)$ and

$$(5.12) \quad 0 \leq n(x, s, c_{i+1}) - n(x, s, d_i) \leq 1.$$

Moreover, $U_{[M, M+L]}^f(x, s)$ is a union of intervals of length $\geq \frac{\inf_{\mathbb{T}} f}{2} \geq \frac{c_{f,g}}{2}$ and similarly $U_{[M, M+L]}^g(y, r)$ is a union of intervals of length $\geq \frac{\inf_{\mathbb{T}} g}{2} \geq \frac{c_{f,g}}{2}$. Therefore the number of intervals in $U_{[M, M+L]}^f(x, s)$ and $U_{[M, M+L]}^g(y, r)$ is bounded by $\frac{2L}{c_{f,g}}$. So $v \leq \frac{4}{c_{f,g}}L$. Define

$$a(t) := t + f^{(n(x, s, t))}(x') - f^{(n(x, s, t))}(x) - p.$$

⁶Use $M' \geq N' \geq c_{f,g}^{-4} + 1$ which implies $2M' \geq M' + c_{f,g}^{-1} + 1 \geq n(x, s, M)$, then property (F1) and finally that $M \geq 6$.

Notice that $a(t) = t + R_i$ on each (c_i, d_i) (since $n(x, s, \cdot)$ is constant on (c_i, d_i)) and (cf. (5.12))

$$\begin{aligned} |R_{i+1} - R_i| &= \\ \left| \left(f^{(n(x, s, c_{i+1}))}(x') - f^{(n(x, s, c_{i+1}))}(x) \right) - \left(f^{(n(x, s, d_i))}(x') - f^{(n(x, s, d_i))}(x) \right) \right| &\leq \\ |f(T^{n(x, s, d_i)}x) - f(T^{n(x, s, d_i)}x')| &\leq \epsilon', \end{aligned}$$

the last inequality by (F3), (5.8) and the definition of κ' . Using (F3) and (5.8) again, in view of (5.10) and (5.5), we have

$$\begin{aligned} |R_1| &= |f^{(n(x, s, c_1))}(x') - f^{(n(x, s, c_1))}(x) - p| \leq \\ n(x, s, c_1)\kappa'^2 + |p| &\leq 2M'\kappa'^2 \leq \frac{4\kappa'^2}{\int f d\mu} M < \epsilon L. \end{aligned}$$

So $(a, U, \frac{c_{f,g}}{4})$ is ϵ -good. Let us also define $b(t) := t + g^{(m(y, r, t))}(y') - g^{(m(y, r, t))}(y)$. We will show that for $t \in U$, we have

$$(5.13) \quad d_1^f(T_t^f(x, s), T_{a(t)+p}^f(x', s')) < \epsilon/200 \text{ and } d_2^g(S_t^g(y, r), S_{b(t)}^g(y', r')) < \epsilon/200.$$

Let us show the first inequality, the proof of the second being analogous.⁷ By the definition of special flow and since $t \in U$, we have

$$(5.14) \quad f^{(n(x, s, t))}(x) + \frac{\epsilon}{100} \leq t + s < f^{(n(x, s, t)+1)}(x) - \frac{\epsilon}{100}.$$

Notice that

$$(5.15) \quad f^{(n(x, s, t))}(x') < a(t) + p + s' < f^{(n(x, s, t)+1)}(x').$$

Indeed, by (5.14) we have

$$a(t) + p + s' = t + s - f^{(n(x, s, t))}(x) + (s' - s) + f^{(n(x, s, t))}(x') \geq \frac{\epsilon}{100} - \delta + f^{(n(x, s, t))}(x') \geq f^{(n(x, s, t))}(x'),$$

since $\delta < \frac{\epsilon}{100}$. Similarly, by the cocycle identity⁸ and (5.14)

$$\begin{aligned} a(t) + p + s' &= t + s - f^{(n(x, s, t))}(x) + (s' - s) + f^{(n(x, s, t))}(x') < \\ f(T^{n(x, s, t)}x) + f^{(n(x, s, t))}(x') - \frac{\epsilon}{100} + \delta &< f^{(n(x, s, t)+1)}(x'), \end{aligned}$$

the last inequality since $|f(T^{n(x, s, t)}x) - f(T^{n(x, s, t)}x')| < \frac{\epsilon}{1000}$ (see (F3)) and $\delta < \frac{\epsilon}{1000}$ what completes the proof of (5.15). Therefore, by the definition of special flow and (5.15), we have

$$T_t^f(x, s) = (T^{n(x, s, t)}x, t + s - f^{(n(x, s, t))}(x))$$

and

$$T_{a(t)+p}^f(x', s') = (T^{n(x, s, t)}x', a(t) + p + s' - f^{(n(x, s, t))}(x')).$$

Hence, the first inequality in (5.13) follows because of the definition of $a(t)$, since d_1^f is the product metric, T is an isometry and $\delta < \epsilon/1000$. Analogously, we show the second inequality which completes the proof of (5.13).

Notice that by (5.13), for $t \in U$, by the triangle inequality⁹, we have

$$(5.16) \quad d_2^g(S_{b(t)}^g(y', r'), \partial Y^g) \geq d_2^g(S_t^g(y, r), \partial Y^g) - d_2^g(S_{b(t)}^g(y', r'), S_t^g(y, r)) \geq \epsilon/200.$$

⁷In the calculation below $s = s'$ but, in general, $r \neq r'$.

⁸Applied to RHS of (5.14).

⁹Here $\partial Y^g := \{(y, 0) : y \in Y\} \cup \{(y, g(y)) : y \in Y\}$.

Therefore, by the triangle inequality, for $t \in U$

$$\begin{aligned} d_2^g(S_t^g(y, r), S_{a(t)}^g(y', r')) &\leq \\ d_2^g(S_t^g(y, r), S_{b(t)}^g(y', r')) + d_2^g(S_{b(t)}^g(y', r'), S_{a(t)-b(t)}^g(S_{b(t)}^g(y', r'))) &\leq \\ \epsilon/200 + d_2^g(S_{b(t)}^g(y', r'), S_{a(t)-b(t)}^g(S_{b(t)}^g(y', r'))) &\leq \end{aligned}$$

So to finish the proof of our proposition, by (5.16), it is enough to show that

$$|b(t) - a(t)| < \epsilon/300 \text{ for } t \in U.$$

which by the definition of $a(t)$, $b(t)$, follows by showing

$$|(g^{(m(y,r,t))}(y') - g^{(m(y,r,t))}(y)) - (f^{(n(x,s,t))}(x') - f^{(n(x,s,t))}(x)) - p| < \epsilon/300$$

whenever t is such that $n(x, s, t) \in [M', M' + L'] \cap \mathbb{Z}$ (see (5.8)). But by property (F1) and (5.10),

$$\begin{aligned} |m(y, r, t) - \zeta n(x, s, t)| &= \frac{1}{\int g d\nu} \left| m(y, r, t) \int g d\nu - n(x, s, t) \int f d\mu \right| \leq \\ 2t^{1/3} &\leq 4M^{1/3} < M^{1/2}, \end{aligned}$$

so by (F4), the above follows by

$$|(g^{([\zeta w])}(y') - g^{([\zeta w])}(y)) - (f^{(w)}(x') - f^{(w)}(x)) - p| < \epsilon/600$$

for $w = n(x, s, t) \in [M', M' + L'] \cap \mathbb{Z}$, which in turn follows from property (F2) since $\epsilon' < \epsilon/1000$. Hence, (3.2) holds and the proof is complete. \square

6. ARNOL'D FLOWS AND BIRKHOFF SUMS ESTIMATES

6.1. Definition of a class of Arnol'd flows. The class of Arnol'd special flows $(R_\alpha)_t^f$ acting on \mathbb{T}^f which we consider consists of special flows over a rotation R_α such that α and f satisfy the following assumptions.

Assumptions on the base rotation. Let $\alpha \in \mathbb{T} \setminus \mathbb{Q}$ and let (q_n) denote the sequence of denominators of α .

Definition 6.1 (the Diophantine condition \mathcal{D}). We say that $\alpha \in \mathcal{D}$ if $\alpha \in \mathbb{T}$ and α satisfies:

(D1) Let $K_\alpha := \{n \in \mathbb{N} : q_{n+1} \leq q_n \log^{7/8} q_n\}$. Then

$$\sum_{i \notin K_\alpha} \log^{-7/8} q_i < +\infty.$$

(D2) There exists a subsequence (n_k) such that

$$q_{n_k+1} \geq q_{n_k} \log(q_{n_k}) \log(\log q_{n_k}), \text{ for } k \geq 1.$$

(D3) For every $n \geq 1$ sufficiently large, we have $q_{n+1} \leq q_n \log^2 q_n$.

Thus, for some constant $D_\alpha > 0$, we have $q_{n+1} \leq D_\alpha q_n \log^2 q_n$ for every $n \in \mathbb{N}$.

Let us remark that condition (D1) first appeared in the work [10] by Fayad and the first author, where it was introduced since it plays a crucial role in proving the SR-property.

Lemma 6.1. *The set \mathcal{D} has full Lebesgue measure in \mathbb{T} .*

Proof. We claim that $(\mathcal{D}2)$, $(\mathcal{D}3)$ are also full (Lebesgue) measure conditions by Khinchine's theorem. Indeed, recall that Khintchine theorem states that, given $\psi : \mathbb{N} \rightarrow \mathbb{R}^+$, the inequality

$$(6.1) \quad \left| \alpha - \frac{p}{q} \right| < \frac{\psi(q)}{q}$$

is satisfied infinitely often for Lebesgue a.e. α if $\sum_q \psi(q) = \infty$, or Lebesgue a.e. α does not satisfy the inequality infinitely often if the series converges. For $(\mathcal{D}2)$, consider $\psi(q) = \frac{1}{q \log q \log \log q}$ and use Legendre theorem to conclude that for the infinitely many solutions p_{n_k}/q_{n_k} of (6.1) we have that q_{n_k} are denominators of α . For $(\mathcal{D}3)$, one can reason analogously using the function $\psi(q) = \frac{1}{q \log^2 q}$. \square

Assumptions on the roof function. We assume that $f : \mathbb{T} \rightarrow \mathbb{R}^+$ is of the form:

$$(6.2) \quad f(x) = A_-(-\log x) + A_+(-\log(1-x)) + g(x),$$

where $g \in C^4(\mathbb{T})$ and

$$A_-, A_+ > 0 \text{ and } A_- \neq A_+.$$

It follows that $f \in C^4(\mathbb{T} \setminus \{0\})$, $f > 0$ and the behavior around 0 is given by the following:

$$(R1) \quad \lim_{x \rightarrow 0^+} \frac{f(x)}{-\log(x)} = A_- \text{ and } \lim_{x \rightarrow 0^+} \frac{f(x)}{-\log(1-x)} = A_+;$$

$$(Rj) \quad \lim_{x \rightarrow 0^+} \frac{\frac{d^j f}{dx^j}(x)}{(-1)^j x^{-j}} = (j-1)!A_- \text{ and } \lim_{x \rightarrow 0^+} \frac{\frac{d^j f}{dx^j}(x)}{(1-x)^{-j}} = (j-1)!A_+ \text{ for } j = 1, 2, 3, 4.$$

6.2. Denjoy-Koksma estimates. In what follows, for simplicity of notation, we assume that $\int_{\mathbb{T}} f d\lambda = 1$. For $x \in \mathbb{T}$ and $n \in \mathbb{N}$ define $i_{n,x} = i(n, x)$ so that $0 \leq i_{n,x} < q_n$ satisfies

$$(6.3) \quad \|x + i_{n,x}\alpha\| = \min_{0 \leq j < q_n} \|x + j\alpha\|.$$

Set also

$$(6.4) \quad B_{n,x} := q_n \|x + i_{n,x}\alpha\| = \min_{0 \leq j < q_n} q_n \|R_\alpha^j x\|.$$

The quantity $B_{n,x}$ will play a crucial role in all the following estimates of the Birkhoff sums f and its derivatives, since these depend on the contribution of the closest visit to the origin of \mathbb{T} , given by $\|x + i_{n,x}\alpha\| = B_{n,x}/q_n$.

Remark 6.1. Remark that we have

$$0 < B_{n,x} < 1 \quad \text{for each } n \geq 0.$$

This can be seen by considering the interval $(-1/q_n, 1/q_n)$ which has length $1/2q_n$, and applying Lemma 2.3, which in particular guarantees that there exists a point of the orbit $\{x + j\alpha, 0 \leq j < q_n\}$ which enters it and thus gives $B_{n,x} < 1$.

6.3. Estimates of Birkhoff sums at return times. The following lemma, which provides estimates for Birkhoff sums of f and its derivatives at special times given by the denominators of α (which corresponds dynamically to closest returns of the orbit to zero), is a consequence of the Denjoy-Koksma inequality.

Lemma 6.2 (Special time estimates). *There exists $C(f) > 0$ such that for every $n \in \mathbb{N}$ and $x \in \mathbb{T}$, we have:*

$$(6.5) \quad \left| f^{(q_n)}(x) - q_n \right| < C(f)(\log q_n + |\log B_{n,x}|),$$

$$(6.6) \quad \left| f'^{(q_n)}(x) - (A_- - A_+)q_n \log q_n \right| < C(f)q_n \left(1 + \frac{1}{B_{n,x}} \right),$$

$$(6.7) \quad \left| f'''(q_n)(x) - f''(x + i\alpha) \right| \leq C(f)q_n^2,$$

$$(6.8) \quad \left| f''''(q_n)(x) - f'''(x + i\alpha) \right| \leq C(f)q_n^3,$$

$$(6.9) \quad \left| f'''''(q_n)(x) - f''''(x + i\alpha) \right| \leq C(f)q_n^4$$

for all $i = 0, \dots, q_n - 1$.

Since the function f and its derivatives do not have bounded variation, the proof is based on defining suitable *truncations* to which Denjoy-Koksma inequality can be applied.

Proof. For given $n \in \mathbb{N}$, let us consider the truncation $\bar{f}_n(x) := \mathbb{1}_{[-\frac{1}{4q_n}, \frac{1}{4q_n}]}(x)f(x)$. Then, since (by Lemma 2.3) there is at most one point (namely $x + i_{n,x}\alpha$) of the form $x + j\alpha$, $j = 0, \dots, q_n - 1$, in the interval $[-\frac{1}{4q_n}, \frac{1}{4q_n}]$, we have

$$(6.10) \quad \left| \bar{f}_n^{(q_n)}(x) - f^{(q_n)}(x) \right| = \begin{cases} 0 & \text{if } x + i\alpha \in [-\frac{1}{4q_n}, \frac{1}{4q_n}], \\ f(x + i\alpha) & \text{otherwise.} \end{cases}$$

Moreover, recalling (6.4) and form of f (see (6.2)),

$$f(x + i_{n,x}\alpha) \leq C_{A_-, A_+, g}(f(\|x + i_{n,x}\alpha\|)) = C_{A_-, A_+, g}\left(f\left(\frac{B_{n,x}}{q_n}\right)\right) \leq C_f(\log q_n + |\log B_{n,x}|).$$

Besides, by the Denjoy-Koksma inequality for \bar{f}_n , we get

$$\left| \bar{f}_n^{(q_n)}(x) - q_n \int_{\mathbb{T}} \bar{f}_n d\lambda \right| < 2\text{Var}(\bar{f}_n).$$

Now,

$$\begin{aligned} \left| \int_{\mathbb{T}} \bar{f}_n d\lambda - 1 \right| &= \left| \int_{\mathbb{T}} \bar{f}_n d\lambda - \int_{\mathbb{T}} f d\lambda \right| = \\ &= A_- \int_0^{1/(4q_n)} \log t dt + A_+ \int_{1-1/(4q_n)}^1 \log(1-t) dt + O(1/q_n) = \\ &= A_- \int_0^{1/(4q_n)} (\log t + 1) dt + A_+ \int_{1-1/(4q_n)}^1 (\log(1-t) - 1) dt + O(1/q_n) = \\ &= A_-(t \log t)|_0^{1/(4q_n)} + A_+((1-t) \log(1-t))|_{1-1/(4q_n)}^1 + O(1/q_n) = O\left(\frac{\log q_n}{q_n}\right). \end{aligned}$$

Since $\text{Var}(\bar{f}_n) = O(\log q_n)$, this completes the proof of (6.5).

We have $f'(x) = \frac{-A_-}{x} + \frac{A_+}{1-x} + g'(x)$. Now, (6.10) holds with f replaced by f' , so

$$\left| \bar{f}_n^{(q_n)}(x) - f'^{(q_n)}(x) \right| \leq \left| f' \left(\frac{B_{n,x}}{q_n} \right) \right| = O\left(\frac{q_n}{B_{n,x}}\right).$$

Moreover, by the Denjoy-Koksma inequality for \bar{f}'_n , we get

$$\left| \bar{f}'_n^{(q_n)}(x) - q_n \int_{\mathbb{T}} \bar{f}'_n d\lambda \right| < 2\text{Var}(\bar{f}'_n).$$

Then $\text{Var} \bar{f}'_n = O(q_n)$ and

$$\begin{aligned} q_n \int_{\mathbb{T}} \bar{f}'_n d\lambda &= q_n \int_{1/(4q_n)}^{1-1/(4q_n)} f' d\lambda = q_n \left(f \left(1 - \frac{1}{4q_n} \right) - f \left(\frac{1}{4q_n} \right) \right) = \\ &= q_n(-A_+ \log q_n + A_- \log q_n + O(1)) = (A_- - A_+)q_n \log q_n + O(q_n). \end{aligned}$$

Hence, (6.6) follows.

We now study the second derivative: $f''(x) = \frac{A_-}{x^2} + \frac{A_+}{(1-x)^2} + g''(x)$. Again, (6.10) holds with f replaced by f'' . This we can write as

$$f''^{(q_n)}(x) - \mathbb{1}_{\left[\frac{-1}{4q_n}, \frac{1}{4q_n}\right]}(x + i\alpha) f''(x + i\alpha) = \bar{f}_n''^{(q_n)}(x).$$

Hence, by the Denjoy-Koksma inequality (for f_n''), we obtain

$$\left| f''^{(q_n)}(x) - \mathbb{1}_{\left[\frac{-1}{4q_n}, \frac{1}{4q_n}\right]}(x + i\alpha) f''(x + i\alpha) \right| \leq q_n \int_{\mathbb{T}} \bar{f}_n'' d\lambda + 2\text{Var}(\bar{f}_n'').$$

Now, $\text{Var}(\bar{f}_n'') = O(q_n^2)$ and

$$q_n \int_{\mathbb{T}} \bar{f}_n'' d\lambda = q_n \left(f' \left(1 - \frac{1}{4q_n} \right) - f' \left(\frac{1}{4q_n} \right) \right) = O(q_n^2).$$

It follows that

$$\left| f''^{(q_n)}(x) - \mathbb{1}_{\left[\frac{-1}{4q_n}, \frac{1}{4q_n}\right]}(x + i\alpha) f''(x + i\alpha) \right| = O(q_n^2).$$

Therefore, if $x + i\alpha \in \left[\frac{-1}{4q_n}, \frac{1}{4q_n}\right]$ then (6.7) follows, while otherwise $f''(\|x + i\alpha\|) = O(q_n^2)$, so (6.7) holds again.

The proofs of remaining inequalities follow the same lines. \square

6.4. Estimates on Birkhoff sums of f . In this subsection, we prove the following estimates on $f^{(n)}$.

Lemma 6.3 (Birkhoff sums of f). *There exists $T_0 > 0$ such that for every $T \geq T_0$ and $x \in \mathbb{T}$ satisfying*

$$\{x + j\alpha : 0 \leq j \leq T\} \cap \left[-\frac{1}{2T \log^4 T}, \frac{1}{2T \log^4 T} \right],$$

we have

$$|f^{(n)}(x) - n| < T^{1/5} \text{ for all } n \in [0, T] \cap \mathbb{Z}.$$

Proof. Let $n = \sum_{j=1}^k b_j q_j$, where $0 \leq b_j \leq \frac{q_{j+1}}{q_j}$ with $b_k > 0$, be the Ostrowski expansion of n . Then

$$f^{(n)}(x) = \sum_{j=1}^k f^{(b_j q_j)} \left(x + \left(\sum_{i=0}^{j-1} b_i q_i \right) \alpha \right) = \sum_{j=1}^k \left(\sum_{s=0}^{b_j-1} f^{(q_j)}(z_{j,s}) \right),$$

where $z_{j,s} = x + \left(\sum_{i=0}^{j-1} b_i q_i \right) \alpha + s q_j \alpha$. Then, by the estimates given by Lemma 6.2 (see in particular (6.5)), for $j \in \{1, \dots, k\}$, we have

$$|f^{(q_j)}(z_{j,s}) - q_j| \leq C(f)(\log q_j - \log B_{j,z_{j,s}}).$$

Hence

$$\left| f^{(n)}(x) - n \right| = \left| \sum_{j=1}^k \left(\sum_{s=0}^{b_j-1} (f^{(q_j)}(z_{j,s}) - q_j) \right) \right| \leq C(f) \sum_{j=1}^k \sum_{s=0}^{b_j-1} \log \frac{q_j}{B_{j,z_{j,s}}}.$$

Using our assumption, we have

$$\frac{q_j}{B_{j,z_{j,s}}} = \frac{1}{\|z_{j,s} + i(j, z_{j,s})\alpha\|} \leq 2T \log^4 T,$$

(we will determine T_0 later), where $i = i(z_{j,s}, j)$. Therefore,

$$\left| f^{(n)}(x) - n \right| \leq C(f) \left(\sum_{j=1}^k b_j \right) \log(2T \log^4 T) \leq 2C(f)k \left(\max_{1 \leq j \leq k} b_j \right) \log T$$

if T is sufficiently large. The latter expression is bounded from above by $C \log^4 T$ (which in turn is bounded by $T^{1/5}$) for a certain constant $C > 0$ since, given that the sequence of denominators grows exponentially fast, $k \leq C_1(\alpha) \log q_k \leq C_1(\alpha) \log T$ and since by (D3) (of Definition 6.1), $b_j \leq \frac{q_{j+1}}{q_j} \leq C(\alpha) \log^2 q_j$ and $q_k \leq T$, we have

$$\max_{1 \leq j \leq k} b_j \leq C(\alpha) \log^2 q_k \leq C(\alpha) \log^2 T.$$

Combining these inequalities, this concludes the proof. \square

6.5. Estimates on Birkhoff sums of f' . The Birkhoff sums of f' are the most delicate to control, since f' is *not* integrable and we need very precise control on the growth in order to later exploit cancelations between different shearing rates. Let us first state the two estimates on Birkhoff sums of f' which will be used in the rest of the paper (Lemma 6.4 and Lemma 6.5).

The first estimate (in Lemma 6.4) provides a fine control of $f'^{(r)}$ (of order $r \log r$ with optimal control on the constants) as long as one assumes that the points in the orbit of x of length r stay sufficiently far from the singularity. Estimates similar to Lemma 6.4 but in the more general context of interval exchange transformations were proved by the third author in [37] (see Proposition 3.4 in [37], as well as its generalization by Ravotti in [32] and Proposition 4.4 in [21]) and inspired in turn by the work of Kochergin on rotations, see e.g. [24].

Given a constant $M > 0$, for arbitrary $n \geq 0$, define

$$(6.11) \quad \Sigma_n(M) := \bigcup_{i=0}^{Mq_{n+1}} R_\alpha^{-i} \left(\left[-\frac{1}{q_n \log^{7/8} q_n}, \frac{1}{q_n \log^{7/8} q_n} \right] \right).$$

The points in $\Sigma_n(M)$ are those whose orbit get close to the singularity and that should be avoided to have the following estimate (which is proved later in this section).

Lemma 6.4 (Control of Birkhoff sums of f' far from singularities). *Fix any $M > 1$. There exists $\epsilon_0 = \epsilon_0(f)$ such that for every $0 < \epsilon < \epsilon_0$ there exists $N \ni n_0 = n_0(\epsilon)$ such that for all $n \geq n_0$, if $\epsilon^4 q_n \leq r \leq Mq_{n+1}$ and $x \notin \Sigma_n(M)$, we have*

$$(6.12) \quad ((A_- - A_+) - \epsilon^2)r \log r \leq f'^{(r)}(x) \leq ((A_- - A_+) + \epsilon^2)r \log r.$$

Moreover, for every $r \in [0, \epsilon^4 q_n] \cap \mathbb{Z}$, we have

$$(6.13) \quad |f'^{(r)}(x)| < \epsilon^2 q_n \log q_n.$$

The other estimate on the Birkhoff sums of f' which we need later (and is also proved later in this section) is the following.

Lemma 6.5 (Control of Birkhoff sums of f' on good time scales). *Assume that $M > 1$. There exists n_1 such that for every $n \in \mathbb{N}$, $n \geq n_1$, any $x \notin \Sigma_n(M)$ and any integers T and r such that*

$$q_n \log q_n \leq T < q_{n+1}, \quad 0 \leq r < \frac{B_{n,x}T}{2},$$

there exists a constant $C'(f)$ such that we have

$$(6.14) \quad |f'^{(r)}(x) - (A_- - A_+)r \log q_n| \leq C'(f)T.$$

Moreover, there exists $\epsilon_0 = \epsilon_0(f)$ such that for any positive $\epsilon < \epsilon_0$, any x and T as above and every pair of integers $0 \leq r, s < T$ such that in addition $|r - s| \leq \epsilon^3 B_{n,x} T$, we have

$$(6.15) \quad |f'^{(r)}(x) - f'^{(s)}(x) - (A_- - A_+)(r - s) \log q_n| < \epsilon^2 T.$$

Both Lemma 6.4 and Lemma 6.5 will be deduced from the following estimate of Birkhoff sums of f' (which is written in a form which indeed allows one to deduce both previously stated lemmas).

Lemma 6.6 (Resonant control of Birkhoff sums of f'). *There exists a constant $C_f > 0$ such that for every $\delta > 0$ there exists $J = J(\delta) \in \mathbb{N}$ such that, for all integers $r \geq q_J$, if $j_r \geq J$ denotes the unique integer such that $q_{j_r} \leq r \leq q_{j_r+1}$, we have*

$$(6.16) \quad \left| f'^{(r)}(x) - (A_- - A_+)r \log q_{j_r} \right| \leq C_f (r + \delta q_{j_r} \log q_{j_r} + \text{Res}(x, r)),$$

for all $x \in \mathbb{T}$, where

$$\text{Res}(x, r) := \min \left\{ \frac{r}{q_{j_r}} \max_{0 \leq i < r} \frac{1}{\|x + i\alpha\|}, \max_{0 \leq i < r} \frac{1}{\|x + i\alpha\|} + 2q_{j_r+1} \log \left(\frac{r}{q_{j_r}} \right) \right\}.$$

Furthermore, there exists a constant $\delta_\alpha > 0$ (depending on α only) such that the estimate (6.16) with $\delta = \delta_\alpha$ holds for every $r \in \mathbb{N}$.

Let us first show how to deduce the proofs of Lemma 6.4 and Lemma 6.5 from Lemma 6.6, we will then prove Lemma 6.6. We will use several times the following remark.

Remark 6.2. For any $r > 0$, let j_r be such that $q_{j_r} \leq r < q_{j_r+1}$ (so that the Ostrowski expansion of r is a sum over j_r terms). Then, since by the assumption (D3) (of Definition 6.1) on the rotation $r \leq q_{j_r+1} \leq q_{j_r} \log^2 q_{j_r}$ for any large r , we have that

$$\lim_{r \rightarrow \infty} \frac{|\log q_{j_r} - \log r|}{\log r} = \lim_{r \rightarrow \infty} \frac{\log \frac{r}{q_{j_r}}}{\log r} \leq \lim_{r \rightarrow \infty} \frac{\log(\log^2 q_{j_r})}{\log q_{j_r}} = 0.$$

Proof of Lemma 6.4. Consider the Ostrowski expansion $r = \sum_{j=0}^{j_r} b_j q_j$, where $b_j \leq a_j$ and $b_{j_r} \geq 1$. We claim that the first part of the lemma will be proved once we show that there exists n_0 such that for $n \geq n_0$ and r as in the assumptions, we have

$$(6.17) \quad \left| f'^{(r)}(x) - (A_- - A_+)r \log q_{j_r} \right| \leq \frac{\epsilon^2}{2} (r \log r).$$

Remark first that, since by assumption $r \geq \epsilon^4 q_n$, we can guarantee that r is sufficiently large if n is sufficiently large. Thus, by Remark 6.2, there is n_0 such that for $n \geq n_0$ and r as in the assumptions

$$|(A_- - A_+)r \log q_{j_r} - (A_- - A_+)r \log r| \leq \frac{\epsilon^2}{2} (r \log r).$$

Thus, once we prove (6.17), combining it with the above equation we get the first part of the lemma (namely (6.12)).

Let us now prove (6.17). By Lemma 6.6 for $\delta = \epsilon^2/6C_f$, the trivial bounds $q_{j_r} \leq r$ and the assumption $r \geq \epsilon^4 q_n$, we get the estimate

$$(6.18) \quad \left| f'^{(r)}(x) - (A_- - A_+)r \log q_{j_r} \right| \leq C_f \left(r + \frac{\epsilon^2}{6C_f} r \log r + \text{Res}(x, r) \right)$$

$$(6.19) \quad \leq \left(\frac{C_f}{\log(q_n \epsilon^4)} + \frac{\epsilon^2}{6} + \frac{C_f \text{Res}(x, r)}{r \log r} \right) r \log r.$$

We now want to show that each of the three terms in parentheses can be bounded by $\frac{\epsilon^2}{2}$. Since for n sufficiently large the first term in the parenthesis is less than $\epsilon^2/6$, we only have to show that $\text{Res}(x, r)/r \log r \leq \epsilon^2/6C_f$. To see this, remark first that, by the assumptions $r \leq Mq_{n+1}$ and $x \notin \Sigma_n(M)$, we have that

$$\min_{0 \leq i < r} \|x + i\alpha\| \geq 1/(q_n \log^{7/8} q_n).$$

Then, consider two cases:

- if $q_{j_r} \geq q_n$, so that $q_n/q_{j_r} \leq 1$, we use the first estimate for $\text{Res}(x, r)$ given by Lemma 6.6 (together with the assumption $r \geq \epsilon^4 q_n$), which gives

$$\frac{\text{Res}(x, r)}{r \log r} \leq \frac{\frac{r}{q_{j_r}} \left(q_n \log^{7/8} q_n \right)}{r \log r} \leq \frac{\frac{q_n}{q_{j_r}} \log^{7/8} q_n}{\log(\epsilon^4 q_n)} \leq \frac{\log^{7/8} q_n}{\log(\epsilon^4 q_n)} < \frac{\epsilon^2}{6C_f}$$

if $n \geq n_0$ as long as n_0 is large enough.

- Otherwise, if $q_{j_r} < q_n$ (and hence $q_{j_r+1} \leq q_n$), we use the other estimate for $\text{Res}(x, r)$ given by Lemma 6.6. Using also that, by Remark 6.2, $\log \frac{r}{q_{j_r}} \leq \epsilon^7 \log r$ if r is large and the assumption $r \geq \epsilon^4 q_n$, it gives

$$\frac{\text{Res}(x, r)}{r \log r} \leq \frac{q_n \left(\log^{7/8} q_n \right) + 2q_{j_r+1} \log(r/q_{j_r})}{r \log r} \leq \frac{\log^{7/8} q_n}{\epsilon^4 \log(\epsilon^4 q_n)} + \frac{2q_n \epsilon^7 \log r}{\epsilon^4 q_n \log r},$$

which can be made less than $\frac{\epsilon^3}{6C_f}$ if $n \geq n_0$ with n_0 large enough and $\epsilon < \epsilon_0$, where

$$\epsilon_0 := \min \left\{ \frac{1}{24C_f}, 1 \right\}.$$

This concludes the proof of (6.17) and hence, by the initial claim, of the first part of the lemma.

To prove the second part (namely (6.13)), let us apply the final part of Lemma 6.6 (e.g. the estimate (6.16) with $\delta = \delta_\alpha$, which is valid for any r), which gives

$$(6.20) \quad |f^{(r)}(x)| \leq (A_- - A_+)r \log q_{j_r} + C_f(r + \delta_\alpha q_{j_r} \log q_{j_r} + \text{Res}(x, r)).$$

To estimate $\text{Res}(x, r)$, we use the second estimate given by Lemma 6.6. Since $x \notin \Sigma(M)$ with $M > 1$, using also assumption (D3) of Definition 6.1 on α and $q_{j_r+1} \leq q_n$ (since $q_{j_r} \leq \epsilon^4 q_n < q_n$), we get

$$\text{Res}(x, r) \leq \max_{0 \leq \ell < r} \frac{1}{\|x + \ell\alpha\|} + 2q_{j_r+1} \log(D_\alpha \log^2 q_{j_r}) \leq q_n (\log q_n)^{7/8} + 4q_n \log(D_\alpha^{1/2} \log q_n),$$

which is less than $\epsilon^4 q_n \log q_n$ if $n \geq n_0$ for n_0 is sufficiently large. Thus, combining this with the initial estimate (6.20) and using that $q_{j_r} \leq r \leq \epsilon^4 q_n$, if $\epsilon_0 < 1$ is chosen so that $\epsilon_0^2 4(A_- - A_+ + C_f(2 + \delta_\alpha)) < 1$ (recalling also the assumption on T), gives that, for any $\epsilon < \epsilon_0$,

$$|f^{(r)}(x)| \leq \epsilon^4 q_n \log q_n ((A_- - A_+) + C_f(2 + \delta_\alpha)) < \epsilon^2 q_n \log q_n < \epsilon^2 T.$$

This concludes also the proof of the second part. \square

Let us now prove Lemma 6.5 using again Lemma 6.6.

Proof of Lemma 6.5. Recall first that $B_{n,x} < 1$ (see Remark 6.1), hence the assumptions $r < B_{n,x}T/2$ and $T < q_{n+1}$ imply in particular that $r < T$ and that $j_r \leq n$ (since

$q_{j_r} \leq r < T < q_{n+1}$). Applying the last part of Lemma 6.6, i.e. the estimate with $\delta = \delta_\alpha$ which holds for any r , (and using the assumption $T > q_n \log q_n$) we get

$$\begin{aligned} |f^{(r)}(x) - (A_- - A_+)r \log q_{j_r}| &< C_f(r + \delta_\alpha q_{j_r} \log q_{j_r} + \text{Res}(x, r)) \\ &\leq C_f(T + \delta_\alpha q_n \log q_n + \text{Res}(x, r)) \leq C_f(T(1 + \delta_\alpha) + \text{Res}(x, r)). \end{aligned}$$

We will now show that $\text{Res}(x, r) \leq T$, so that the RHS is bounded by $C_f(2 + \delta_\alpha)T$. If $r \geq q_n$, since we then have $q_{j_r} = q_n$ (as we already observed above that $j_r \leq n$), this estimate will conclude the proof. If, on the other hand, $r < q_n$, we also have $q_{j_r} < q_n$, so that the terms $(A_- - A_+)r \log q_{j_r}$ and $(A_- - A_+)r \log q_n$ are both less than $(A_- - A_+)q_n \log q_n < (A_- - A_+)T$ (by the assumptions on T). Thus also in this case we conclude the proof, by setting $C'(f) := C_f(2 + \delta_\alpha) + 2(A_- - A_+)$.

We are hence now left to show that $\text{Res}(x, r) \leq T$. To see this, let us first show that, for any $r < B_{n,x}T/2$, we have

$$(6.21) \quad \max_{0 \leq \ell < r} \frac{1}{\|x + \ell\alpha\|} \leq \min \left\{ q_n \log^{7/8} q_n, \frac{2q_n}{B_{n,x}} \right\}.$$

The first bound (by $q_n \log^{7/8} q_n$) follows simply from the assumption $x \notin \Sigma_n(M)$ and $M > 1$ (recall the definition (6.11) of $\Sigma_n(M)$). Let us hence prove the second bound (namely by $2q_n/B_{n,x}$). Let us denote by $m_{x,r} = m(x, r)$ the index $0 \leq m_{x,r} < r$ such that

$$\|x + m_{x,r}\alpha\| = \min\{\|x + \ell\alpha\|, 0 \leq \ell < r\}.$$

Since $m_{x,r} < r$, dividing $m_{x,r}$ by q_n we get that there exists $0 \leq j \leq r/q_n$ such that $0 \leq m_{x,r} - jq_n \leq q_n$. Since the points $(R_{q_n\alpha})^i x$ and $(R_{q_n\alpha})^{i+1} x$ for $0 \leq i < r - 1$ are at most $1/q_{n+1}$ apart (recall (2.5)), using also the assumptions $r < B_{n,x}T/2$ and $T < q_{n+1}$, it follows that

$$\|(x + (m_{x,r} - jq_n)\alpha) - (x + m_{x,r}\alpha)\| \leq \frac{r}{q_n} \frac{1}{q_{n+1}} \leq \frac{B_{n,x}T}{2q_n} \frac{1}{q_{n+1}} \leq \frac{B_{n,x}}{2q_n}.$$

Moreover, by the definition of $B_{n,x}$ (recall (6.4)), we have that $B_{n,x}/q_n \leq \|x + (m_{x,r} - jq_n)\alpha\|$ and by Lemma 2.3 (since $r > q_n$) that $\|x + m_{x,r}\alpha\| < 1/q_n$. It follows that $x' := x + (m_{x,r} - jq_n)\alpha$ and $x'' := x + m_{x,r}\alpha$ either both belong to the interval $\left[\frac{B_{n,x}}{2q_n}, \frac{3}{2q_n}\right]$ or both to the interval $\left[1 - \frac{3}{2q_n}, 1 - \frac{B_{n,x}}{2q_n}\right]$, so that (for n large) we either have that $\|x'\| = x'$ and $\|x''\| = x''$, or we have that $\|x'\| = 1 - x'$ and $\|x''\| = 1 - x''$. Thus, from the previous equation (and the definition of $B_{n,x}$), we also deduce that

$$\frac{B_{n,x}}{q_n} \leq \|x + (m_{x,r} - jq_n)\alpha\| \leq \|x + m_{x,r}\alpha\| + \frac{B_{n,x}}{2q_n}.$$

Recalling the definition of $m_{x,r}$, this gives that $\min\{\|x + \ell\alpha\|, 0 \leq \ell < r\} \geq \frac{B_{n,x}}{2q_n}$, which concludes the proof of (6.21).

Let us go back to show the estimate $\text{Res}(x, r) < T$. We consider separately two cases:

- Assume first that $q_{j_r} < q_n$. Remark that this means that $q_{j_r+1} \leq q_n$ and hence $r \leq q_n$. Thus, by this remark and since by Remark 6.2 for any $0 < \delta < 1$ we have $\log(r/q_{j_r}) \leq \delta \log r/4$ for r large (and recalling that $T > q_n \log q_n$), we get

$$\text{Res}(x, r) \leq q_n \log q_n^{7/8} + 2q_{j_r+1} \log \left(\frac{r}{q_{j_r}} \right) \leq q_n \log q_n \left(\frac{1}{\log^{1/8} q_n} + \frac{\delta}{2} \right) < \delta T$$

for n sufficiently large.

- Otherwise, if $q_{j_r} \geq q_n$, we have $q_{j_r} = q_n$ (since by the assumptions $j_r \leq n$). In this case using (6.21) and the assumption $r < B_{n,x}T/2$, we can then estimate

$$\text{Res}(x, r) \leq \frac{r}{q_n} \max_{0 \leq \ell < r} \frac{1}{\|x + \ell\alpha\|} \leq \frac{B_{n,x}T}{2q_n} \frac{2q_n}{B_{n,x}} = T.$$

This concludes also the first part of the lemma.

For the second part of the lemma, assuming without loss of generality that $r > s$ and using the cocycle property of Birkoff sums, we have that $|f^{(r)}(x) - f^{(s)}(x)| = |f^{(r-s)}(z)|$ for $z := R_\alpha^s(x)$. We claim that, by Lemma 6.6, for any $\epsilon > 0$ there exists $n_0 > 0$ such that for all $n \geq n_0$, we have that, for any r, s as in the assumptions,

$$(6.22) \quad |f^{(r)}(x) - f^{(s)}(x) - (A_- - A_+)(r - s) \log q_{j_{r-s}}| < C_f(r - s) + \frac{\epsilon^2}{6} q_n \log q_n + C_f \text{Res}(z, r - s).$$

To see this, pick $N = N(\epsilon, f)$ such that,

$$(6.23) \quad q_{n-N} \leq \frac{1}{\sqrt{2}^N} q_n < \left(\min \left\{ \frac{\epsilon^2}{6\delta_\alpha}, \frac{\epsilon^2}{4(A_- - A_+)} \right\} \right) q_n,$$

where δ_α is the constant in the last part of Lemma 6.6 (the second bound for N will be used only later in the proof). Applying the first part of Lemma 6.6 for $\delta = \epsilon^2/(6C_f)$, (since $n - N$ goes to infinity as n does) there exists n_0 such that for all $n \geq n_0$ (6.22) holds as long as $r - s \geq q_{n-N}$ (remark that $q_{j_{r-s}} \leq q_n$ since $r - s < T < q_{n+1}$). On the other hand, if $r - s \leq q_{n-N}$, we can use the final part of Lemma 6.6 to get that

$$|f^{(r)}(x) - f^{(s)}(x) - (A_- - A_+)(r - s) \log q_{j_{r-s}}| < C_f(r - s) + \delta_\alpha q_{j_{r-s}} \log q_{j_{r-s}} + C_f \text{Res}(z, r - s)$$

which also in this case gives (6.23) since by assumption on r, s and choice (6.23) of N we have that $\delta_\alpha q_{j_{r-s}} \log q_{j_{r-s}} \leq \delta_\alpha q_{n-N} \log q_n \leq \frac{\epsilon^2}{6} q_n \log q_n$. This concludes the proof of (6.22) and the claim.

We now want to show that for some $n_1 \geq n_0$ and $\epsilon_0 > 0$ such that each of the three terms in the RHS of (6.22) is less than $(\epsilon^2/6)T$ if $\epsilon < \epsilon_0$ and $n \geq n_1$. Since by assumption $r - s \leq \epsilon^3 T$ (recall that $B_{n,x} < 1$, by Remark 6.1) to control the first term, it is enough to choose $\epsilon < \epsilon_0$, where $\epsilon_0 := 1/(6C_f)$. For the second is enough to recall that by assumption $q_n \log q_n < T$. Finally, to estimate $\text{Res}(z, r - s)$ we use two regimes:

- either $q_{j_{r-s}} \geq q_n$ and hence $q_{j_{r-s}} = q_n$ (since as we already remarked $r - s < q_{n+1}$), in which case we use the first estimate for $\text{Res}(z, r - s)$ (in Lemma 6.6) combined with (6.21) and the assumption on $(r - s)$ to get

$$\begin{aligned} \text{Res}(z, r - s) &\leq \frac{r - s}{q_{j_{r-s}}} \max_{0 \leq \ell < r-s} \frac{1}{\|z + \ell\alpha\|} \leq \frac{r - s}{q_n} \max_{0 \leq \ell < r} \frac{1}{\|x + \ell\alpha\|} \\ &\leq \frac{\epsilon^3 B_{n,x} T}{q_n} \frac{2q_n}{B_{n,x}} = 2\epsilon^3 T, \end{aligned}$$

which is less than $(\epsilon^2 T)/(6C_f)$ for $\epsilon < \epsilon_0$, given that $\epsilon_0 < 1/(12C_f)$.

- Otherwise, if $q_{j_{r-s}} < q_n$ (and hence $q_{j_{r-s}+1} \leq q_n$), we use the other estimate (and assumption (D3) of Definition 6.1 on α) to get

$$\begin{aligned} \text{Res}(z, r - s) &\leq \max_{0 \leq \ell < r-s} \frac{1}{\|z + \ell\alpha\|} + 2q_{j_{r-s}+1} \log(D_\alpha \log^2 q_{j_{r-s}}) \\ &\leq q_n (\log q_n)^{7/8} + 4q_n \log(D_\alpha^{1/2} \log q_n), \end{aligned}$$

which is less than $\epsilon^3 q_n \log q_n$ if $n \geq n_0$ for n_0 is sufficiently large and hence less than $\frac{\epsilon^2}{6C_f} T$ if $\epsilon < \epsilon_0$.

Thus, we showed that

$$|f'^{(r)}(x) - f'^{(s)}(x) - (A_- - A_+)(r - s) \log q_{j_{r-s}}| \leq \epsilon^2 \frac{T}{2}.$$

In order to conclude, it is now enough to show that

$$(6.24) \quad |(A_- - A_+)(r - s) \log q_{j_{r-s}} - (A_- - A_+)(r - s) \log q_n| \leq \frac{\epsilon^2}{2} T.$$

If $r - s \geq q_n$, we have that $q_{j_{r-s}} = q_n$ (since we also know that $r - s < T < q_{n+1}$) and hence there is nothing to prove. Thus we can assume that $r - s < q_n$. Recall the choice of N made earlier (see (6.23)) and consider again two cases.

- If $q_{n-N} \leq r - s < q_n$, by definition of j_{r-s} we have $q_{j_{r-s}} \geq q_{n-N}$. Since we can write $q_n = (\prod_{j=n-N}^{n-1} \frac{q_{j+1}}{q_j}) q_{n-N}$, using also the assumption $(\mathcal{D}3)$ on α (cf. Def. 6.1), we get in this case that

$$\log \frac{q_n}{q_{j_{r-s}}} \leq \log \frac{q_n}{q_{n-N}} \leq \sum_{j=n-N}^{n-1} \log (D_\alpha \log^2 q_j) \leq 2N \log (D_\alpha^{1/2} \log q_n).$$

Thus, since this bound divided by $\log q_n$ goes to zero as n grows to infinity, increasing n_1 if necessary, for $n \geq n_1$ (recalling that by assumption $q_n \log q_n < T$), we have

$$(r - s) \log \left(\frac{q_n}{q_{j_{r-s}}} \right) \leq q_n \left(\frac{\epsilon^2}{2(A_- - A_+)} \log q_n \right) \leq \frac{\epsilon^2 T}{2(A_- - A_+)}.$$

- Otherwise, if $0 \leq r - s \leq q_{n-N}$, we simply have

$$\begin{aligned} |(r - s) (\log q_{j_{r-s}} - \log q_n)| &\leq 2q_{n-N} \log q_n \\ &\leq 2 \frac{\epsilon^2}{4(A_- - A_+)} q_n \log q_n < \frac{\epsilon^2 T}{2(A_- - A_+)} \end{aligned}$$

by the assumption on T and the choice of N .

In all cases, this concludes the proof of (6.24) and hence of the second part of the lemma. \square

The rest of this section will be devoted to the proof of Lemma 6.6.

Proof of Lemma 6.6. As in the proof of Lemma 6.3, consider the Ostrowski expansion of r , which, by the definition of j_r , is given by $r = \sum_{j=1}^{j_r} b_j q_j$, where $b_j \leq a_j \leq \frac{q_{j+1}}{q_j}$ and $b_{j_r} \geq 1$. Correspondingly, we get the following decomposition of Birkhoff sums of f' :

$$f'^{(r)}(x) = \sum_{j=1}^{j_r} \left(\sum_{s=0}^{b_j-1} f'^{(q_j)}(z_{j,s}) \right), \text{ where } \begin{cases} z_{j_r,s} = x + s q_{j_r} \alpha, \\ z_{j,s} = x + \sum_{k=j+1}^{j_r} b_k q_k \alpha + s q_j \alpha \end{cases} \text{ for } 1 \leq j < j_r.$$

Then, by applying the estimates of Lemma 6.2 (in particular (6.6) for the point $z_{j,s}$), we have

$$(6.25) \quad \left| f'^{(r)}(x) - \sum_{j=1}^{j_r} (A^- - A^+) b_j q_j \log q_j \right| \leq C(f) \sum_{j=1}^{j_r} \sum_{s=0}^{b_j-1} q_j \left(1 + \frac{1}{B_{j,z_{j,s}}} \right).$$

Let us first show that for any $\delta > 0$, there exists $J_0 = J_0(\delta)$ such that for $r \geq q_{J_0}$, we have

$$(6.26) \quad \left| \sum_{j=1}^{j_r} (A^- - A^+) b_j q_j \log q_j - (A^- - A^+) r \log q_{j_r} \right| \leq \frac{\delta}{2} (A^+ - A^-) q_{j_r} \log q_{j_r}.$$

and also that there exists a constant $\delta_\alpha > 0$ depending on α such that the estimate above holds for every $r \in \mathbb{N}$ if we take $\delta = 2\delta_\alpha$ (which is needed to prove the final part of the lemma). To prove the latter, it is enough to use the Ostrowski expansion of r to write

$$\left| \sum_{j=1}^{j_r} b_j q_j \log q_j - r \log q_{j_r} \right| = \sum_{j=1}^{j_r-1} b_j q_j \log \frac{q_{j_r}}{q_j} \leq \left(\sum_{j=1}^{j_r-1} q_j \right) \log q_{j_r}$$

(remark that $\log(q_{j_r}/q_j) = 0$ for $j = j_r$ so the last sum does not include the term j_r) and then using that $\sum_{j=1}^{j_r-1} q_j < q_{j_r}$ by definition of Ostrowski expansion.

To now prove (6.26), choose $N = N(\delta)$ such that $q_{j_r-N}/q_{j_r} \leq 1/\sqrt{2}^N \leq \delta/8$. Writing $q_{j_r} = (\prod_{j=j_r-N}^{j_r-1} \frac{q_{j+1}}{q_j}) q_{j_r-N}$ and using the assumption (D3) on α (cf. Def. 6.1), for any $j_r - N \leq j < j_r$, we have

$$\log \frac{q_{j_r}}{q_j} \leq \log \frac{q_{j_r}}{q_{j_r-N}} \leq \sum_{j=j_r-N}^{j_r-1} \log(\log q_j)^2 \leq 2N \log \log q_{j_r}.$$

Thus, using that, by definition of Ostrowski expansion, $r = \sum_{j=1}^{j_r} b_j q_j$ and $\sum_{j < j_r-N} b_j q_j \leq q_{j_r-N}$, we get that

$$\begin{aligned} \left| r \log q_{j_r} - \sum_{j=1}^{j_r} b_j q_j \log q_j \right| &\leq \left| \sum_{j=j_r-N}^{j_r-1} b_j q_j \log \frac{q_{j_r}}{q_j} \right| + 2 \left(\sum_{j < j_r-N} b_j q_j \right) \log q_{j_r} \\ &\leq q_{j_r} (2N \log \log q_{j_r}) + 2q_{j_r-N} \log q_{j_r} \\ &\leq \left(\frac{2N \log \log q_{j_r}}{\log q_{j_r}} + \frac{\delta}{4} \right) q_{j_r} \log q_{j_r}, \end{aligned}$$

where the last inequality follows from the previous choice of N . Thus, for some $J \geq J_0$, we can hence guarantee that, for $r \geq q_J$, the RHS is less than $\delta/2(q_{j_r} \log q_{j_r})$ and this concludes the proof of (6.26).

Let us now estimate the RHS of (6.25). Recall that $q_j/B_{j,z} = 1/||z + i(j, z)\alpha||$ (by the definition of $B_{j,z}$, cf. (6.4)), so that, recalling also the definition of Ostrowski decomposition, we have

$$(6.27) \quad \sum_{j=1}^{j_r} \sum_{s=0}^{b_j-1} q_j \left(1 + \frac{1}{B_{j,z_{j,s}}} \right) \leq r + \sum_{j=1}^{j_r} \sum_{s=0}^{b_j-1} \frac{1}{||x + i(j, z_{j,s})\alpha||}.$$

For each $1 \leq j \leq j_r$, the points of points $x + i(j, z_{j,s})\alpha$ for $s = 0, 1, \dots, b_j - 1$ are distinct (since they belong to disjoint orbits of length q_j) and they are contained in an orbit of R_α of length q_{j+1} (namely the orbit of the point $z_{j+1, b_{j+1}} := z_{j,0} = x + \sum_{k=j+1}^{j_r} b_k q_k \alpha$ if $1 \leq j < j_r$, or of the point x for $j = j_r$). Thus, by the orbit spacing given by Lemma 2.3

(and recalling the definition of $\|\cdot\|$, see Section 2.4), for $1 \leq j \leq j_r$ we can estimate

$$(6.28) \quad \sum_{s=0}^{b_j-1} \frac{1}{\|x + i(j, z_{j,s})\alpha\|} \leq \sum_{s=0}^{b_j-1} \frac{1}{\{x + i(j, z_{j,s})\alpha\}} + \sum_{s=0}^{b_j-1} \frac{1}{1 - \{x + i(j, z_{j,s})\alpha\}} \\ \leq 2 \sum_{s=0}^{b_j-1} \frac{1}{\min_{k=0, \dots, b_j-1} \|x + i(j, z_{j,k})\alpha\| + s \frac{1}{2q_{j+1}}},$$

where moreover, for $1 \leq j < j_r$,

$$\min_{k=0, \dots, b_j-1} \|x + i(j, z_{j,k})\alpha\| \geq \|x + i(j+1, z_{j+1, b_{j+1}})\alpha\|.$$

One has the following estimate on the sum of inverses of arithmetic progression of step $h > 0$ and length $K \in \mathbb{N}$, starting at $x_0 > 0$ (see for example Lemma 5.1 in [24] or Lemma 9 in [37]):

$$(6.29) \quad \sum_{s=1}^K \frac{1}{x_0 + sh} \leq \frac{1}{x_0} + \frac{\log K}{h}.$$

Thus, applying this estimate for each $1 \leq j < j_r$ and including the contribution of the minimum of each arithmetic progression with $1 \leq j < j_r$ (namely $1/\|x + i(j+1, z_{j+1, s})\alpha\|$) with the arithmetic progression level $j+1$ (notice that it is a distinct term since it belongs to a different block of the orbit), we get

$$(6.30) \quad \sum_{j=1}^{j_r-1} \sum_{s=0}^{b_j-1} \frac{1}{\|x + i(j, z_{j,s})\alpha\|} \leq \sum_{j=1}^{j_r-1} 2q_{j+1} \log(b_j + 1).$$

Using assumption (D3) (of Definition 6.1) on the rotation number, $b_j \leq q_{j+1}/q_j \leq D_\alpha \log^2 q_j$. Thus (since $\sum_{j \leq j_r} q_j/q_{j_r}$ is bounded uniformly in r because q_j grow exponentially) there exists a constant $C_\alpha > 0$ such that for any r we have

$$\frac{\sum_{j=1}^{j_r-1} 2q_{j+1} \log(b_j + 1)}{q_{j_r} \log q_{j_r}} \leq \frac{\log(D_\alpha \log^2 q_{j_r} + 1)}{\log q_{j_r}} \sum_{j=1}^{j_r-1} \frac{2q_{j+1}}{q_{j_r}} \leq C_\alpha.$$

Moreover, the above expression can be made less than $\delta/2$ if $j_r \geq J_1$, up to enlarging J_1 if necessary (since the term before the sum goes to zero as j_r grows while the sum over j , as already remarked, is uniformly bounded in r).

Combining this with (6.27) and (6.30), we have thus shown that for $r \geq q_{J_1}$, we have

$$(6.31) \quad \sum_{j=1}^{j_r} \sum_{s=0}^{b_j-1} q_j \left(1 + \frac{1}{B_{j, z_{j,s}}} \right) \leq r + \frac{\delta}{2} (q_{j_r} \log q_{j_r}) + \text{Res}(x, r),$$

where the term $\text{Res}(x, r)$ (where Res stays for *resonant*) is simply the sum relative to $j = j_r$, namely, recalling the definition (6.4) of B and of $i(\cdot, \cdot)$,

$$\text{Res}(x, r) := \sum_{s=0}^{b_{j_r}-1} \frac{q_{j_r}}{B_{j_r, z_{j_r, s}}} = \sum_{s=0}^{b_{j_r}-1} \frac{1}{\|z_{j_r, s} + i(j_r, z_{j_r, s})\alpha\|} = \sum_{s=0}^{b_{j_r}-1} \max_{0 \leq \ell < q_{j_r}} \frac{1}{\|z_{j_r, s} + \ell\alpha\|}.$$

Moreover, we have also shown that the above estimate (6.31) for $\delta = 2C_\alpha$ holds for all r . Thus, the estimate (6.31) combined with (6.25), (6.26) and concludes the proof of both parts of the lemma when setting $\delta_\alpha := C_\alpha + 1$ and $C_f := \max\{C(f), (A_- - A_+)\}$ as long as we prove the estimate on $\text{Res}(x, r)$ stated in the lemma. This follows from two separate estimates (of which one then takes the minimum). On one hand, we can estimate $\text{Res}(x, r)$

by the number of terms times the largest term, i.e. (since the points $z_{j_r,s} + i(j_r, z_{j_r,s})\alpha$ for $s = 0, \dots, b_{j_r}$ are just points in the orbit of x of length r)

$$\text{Res}(x, r) \leq b_{j_r} \max_{s=0, \dots, b_{j_r}-1} \frac{1}{\|z_{j_r,s} + i(j_r, z_{j_r,s})\alpha\|} \leq \frac{r}{q_{j_r}} \max_{0 \leq \ell < r} \frac{1}{\|x + \ell\alpha\|}.$$

Alternatively, we can estimate $\text{Res}(x, r)$ by using (6.28) for $j = j_r$. Using the bound (6.29) for an arithmetic progression of step $1/q_{j_r+1}$ and length $b_{j_r} \leq r/q_{j_r}$ and bounding trivially the minimum term of the progression (as above), we get

$$\text{Res}(x, r) \leq \max_{0 \leq \ell < r} \frac{1}{\|x + \ell\alpha\|} + 2q_{j_r+1} \log \left(\frac{r}{q_{j_r}} \right).$$

Taking the minimum of these two estimates concludes the estimate of $\text{Res}(x, r)$ and hence the proof of the lemma. \square

6.6. Birkhoff sums of f'' , f''' and f'''' .

Lemma 6.7 (Control of Birkhoff sums of f''). *There exists a constant $C'(f) > 0$ such that for every $\epsilon > 0$ we can find $n_0 \in \mathbb{N}$ such that for every $n \geq n_0$ and $x \in \mathbb{T}$ for which $B_{n,x} \geq \epsilon^{1/5}$ and for all $k \in [1, \frac{B_{n,x}^5 q_{n+1}}{6q_n}] \cap \mathbb{Z}$, we have*

$$(6.32) \quad |f''(kq_n)(x) - kf''(x + i\alpha)| \leq C'(f)kq_n^2,$$

$$(6.33) \quad |f'''(kq_n)(x) - kf'''(x + i\alpha)| \leq C'(f)kq_n^3,$$

$$(6.34) \quad |f''''(kq_n)(x) - kf''''(x + i\alpha)| \leq C'(f)kq_n^4,$$

where $i = i(n, x)$.

Proof. For $s = 0, 1, \dots, k-1$, set $z_s = x + sq_n\alpha$. In view of (6.7) (applied to z_s), we obtain

$$f''(kq_n)(x) = \sum_{s=0}^{k-1} f''(q_n)(z_s) = \sum_{s=0}^{k-1} (f''(z_s + i(z_s, n)\alpha) + O(q_n^2)),$$

whence

$$\begin{aligned} |f''(kq_n)(x) - kf''(x + i\alpha)| &= \left| \sum_{s=0}^{k-1} (f''(z_s + i(z_s, n)\alpha) - f(x + i\alpha)) \right| + O(kq_n^2) \leq \\ &\sum_{s=0}^{k-1} |f''(z_s + i(z_s, n)\alpha) - f(x + i\alpha)| + O(kq_n^2). \end{aligned}$$

Note that if neither $x + i\alpha$ nor $z_s + i(n, z_s)\alpha$ belongs to $\left[\frac{-1}{10q_n}, \frac{1}{10q_n}\right]$ then f'' at each of these points is of order $O(q_n^2)$, so the corresponding summand in the latter sum is of the same order. Suppose now that either $x + i\alpha$ or $z_s + i(n, z_s)\alpha$ belongs to $\left[\frac{-1}{10q_n}, \frac{1}{10q_n}\right]$. We claim that in this case $i(n, x) = i(n, z_s)$. Indeed, since $k \leq B_{n,x}^5 q_{n+1}/(6q_n)$, we have

$$\|x - z_s\| = \|sq_n\alpha\| \leq k\|q_n\alpha\| \leq \frac{B_{n,x}^5 q_{n+1}}{6q_n} \|q_n\alpha\| < \frac{B_{n,x}^5}{6q_n} < \frac{1}{6q_n}.$$

Now, if $x + i\alpha \in \left[\frac{-1}{10q_n}, \frac{1}{10q_n}\right]$ then $z_s + i\alpha \in \left[\frac{-1}{4q_n}, \frac{1}{4q_n}\right]$ which means that $i = i(n, z_s)$ and reversing the role of the two points (if necessary), the claim follows.

So we have $i(n, x) = i = i(n, z_s)$ and our next claim is that the points $x + i\alpha$ and $z_s + i\alpha$ are on the same side of the singularity (at 0). Indeed, as $\|x - z_s\| \leq \frac{B_{n,x}^5}{6q_n} = \|x + i\alpha\|/6$, we only need to note that

$$0 \notin \left(x + i\alpha - \frac{\|x + i\alpha\|}{6}, x + i\alpha + \frac{\|x + i\alpha\|}{6} \right)$$

which is obvious. It follows that we can use the mean value theorem to obtain

$$|f''(x + sq_n\alpha + i\alpha) - f''(x + i\alpha)| = f'''(\theta) \|sq_n\alpha\| \leq f'''(\theta) \frac{B_{n,x}^5}{6q_n},$$

where

$$\theta \in \left(x + i\alpha - \frac{B_{n,x}^5}{6q_n}, x + i\alpha + \frac{B_{n,x}^5}{6q_n} \right).$$

But $B_{n,x} > \epsilon^{1/5}$ and $\|x + i\alpha\| = B_{n,x}/q_n$, so $\|\theta\| \geq B_{n,x}/(4q_n)$. Since $f'''(x) = \frac{-2A_-}{x^3} + \frac{2A_+}{(1-x)^3} + g'''(x)$, we have

$$f'''(\theta) \frac{B_{n,x}^5}{6q_n} = O\left(\frac{4q_n}{B_{n,x}}\right)^3 \frac{B_{n,x}^5}{6q_n} = O(B_{n,x}^2 q_n^2) = O(q_n^2)$$

which completes the proof. \square

Lemma 6.8 (Control of Birkhoff sums of higher derivatives). *There exist $C(f), c(f) > 0$ such that for every $n \in \mathbb{N}$, $x \in \mathbb{T}$ satisfying $B_{n,x} < c(f)$ and for $k \in [1, \frac{B_{n,x}q_{n+1}}{6q_n}] \cap \mathbb{Z}$, we have*

$$(6.35) \quad c(f)kf''(x + i(n, x)\alpha) \leq |f''^{(kq_n)}(x)| \leq C(f)kf''(x + i(n, x)\alpha),$$

$$(6.36) \quad c(f)kf'''(x + i(n, x)\alpha) \leq |f'''^{(kq_n)}(x)| \leq C(f)kf'''(x + i(n, x)\alpha),$$

$$(6.37) \quad c(f)kf''''(x + i(n, x)\alpha) \leq |f''''^{(kq_n)}(x)| \leq C(f)kf''''(x + i(n, x)\alpha).$$

Proof. We will show the proof of (6.35), the proof of (6.36) and (6.37) follows the same lines. As in the proof of previous lemma, we have

$$f''^{(kq_n)}(x) = \sum_{s=0}^{k-1} f''^{(q_n)}(z_s) = \sum_{s=0}^{k-1} f''(z_s + i(z_s, n)\alpha) + O(kq_n^2).$$

We will show first that there is a constant $C > 0$ such that

$$(6.38) \quad C^{-1}f''(x + i(n, x)\alpha) \leq f''(z_s + i(n, z_s)\alpha) \leq Cf''(x + i(n, x)\alpha).$$

Indeed:

- If $x + i(n, x)\alpha$ and $z_s + i(n, z_s)\alpha$ are not in $\left[\frac{-1}{10q_n}, \frac{1}{10q_n}\right]$ then both numbers $f''(x + i(n, x)\alpha)$ and $f''(z_s + i(n, z_s)\alpha)$ are of order q_n^2 since $f''(t) = \frac{A_-}{t^2} + \frac{A_+}{(1-t)^2} + g''(t)$, so (6.38) holds for a relevant constant.
- if $x + i(n, x)\alpha \in \left[\frac{-1}{10q_n}, \frac{1}{10q_n}\right]$ or $z_s + i(n, z_s)\alpha \in \left[\frac{-1}{10q_n}, \frac{1}{10q_n}\right]$ then,

$$\|sq_n\alpha\| \leq s\|q_n\alpha\| \leq \frac{B_{n,x}q_{n+1}}{6q_n} \cdot \frac{1}{q_{n+1}} = \frac{\|x + i(n, x)\alpha\|}{6}.$$

Hence, as before, $i(n, z_s) = i(n, x)$ and

$$z_s + i(n, x)\alpha \in \left[x + i(n, x)\alpha - \frac{\|x + i(n, x)\alpha\|}{6}, x + i(n, x)\alpha + \frac{\|x + i(n, x)\alpha\|}{6} \right].$$

For simplicity, for the rest of the proof, we denote $i = i(n, x)$. Then (for relevant constants)

$$C_1 f''(x + i\alpha) \leq f''\left(\frac{5}{6}(x + i\alpha)\right) = f''\left(x + i\alpha - \frac{x + i\alpha}{6}\right) \leq C'_1 f''(z_s + i\alpha)$$

and a lower bound will be obtained in the same way by using $f''(\frac{7}{6}(x + i\alpha))$ instead of $f''(\frac{5}{6}(x + i\alpha))$, whence (6.38) holds.

Now, in view of (6.38), we have

$$f''(kq_n)(x) = \sum_{s=0}^{k-1} f''(z_s + i(z_s, n)\alpha) + O(kq_n^2) \leq Ck f''(x + i\alpha) + O(kq_n^2) < \tilde{C}k f''(x + i\alpha)$$

as $f''(x + i\alpha)$ is *at least* of order q_n^2 . For the lower bound, by (6.38), we have

$$f''(kq_n)(x) \geq C^{-1}k f''(x + i\alpha) + O(kq_n^2)$$

and we want to show that the latter term is $\geq C''k f''(x + i\alpha)$, that is, for some constant $D > 0$ (implicit in $O(kq_n^2)$), we want to have

$$C^{-1}k f''(x + i\alpha) - Dkq_n^2 > C''k f''(x + i\alpha).$$

Hence, we want to have $(C^{-1} - C'')f''(x + i\alpha) > Dq_n^2$ and for that we only need $f''(x + i\alpha) > D'q_n^2$. \square

Remark 6.3. Note that in the course of the above proof, we showed

$$|f''(kq_n)(x)| = O(k f''(x + i(n, x)\alpha) + q_n^2)$$

with no assumption on $x \in \mathbb{T}$. It follows that the RHS inequality in (6.35) requires no assumption on x .

Recall also that we have $\|x + i(n, x)\alpha\| < \frac{1}{q_n}$ for each $x \in \mathbb{T}$, so we can choose n_0 so that $f''(x + i(n, x)\alpha) > 0$ for each $n \geq n_0$ and all $x \in \mathbb{T}$ (as $f''(y)$ is of order $1/\|y\|^2$).

We also have the following lemma:

Lemma 6.9. *There exist $D > 0$, $n_0 \in \mathbb{N}$ such that for every $x \in \mathbb{T}$, $n \in \mathbb{N}$, $n \geq n_0$ and every $w \in \left[q_{n_0}, \max\left(\frac{B_{n,x}q_{n+1}}{4} - q_n, 0\right)\right] \cap \mathbb{Z}$, we have*

$$(6.39) \quad |f''^{(w)}(x)| \leq D \frac{w}{q_n} f''(x + i(n, x)\alpha).$$

Proof. Fix n and $x \in \mathbb{T}$ and let w be arbitrary. Then

$$(6.40) \quad w = kq_n + s, \quad 0 \leq s < q_n.$$

Since $f''(x) = \frac{A_-}{x^2} + \frac{A_+}{(1-x)^2} + g''(x)$, we have $f''^{(w)}(x) \geq -C_1 w$, where $C_1 = \|g''\|_\infty$. By the same token

$$f''^{((k+1)q_n - w)}(x + w\alpha) \geq -C_1((k+1)q_n - w) \geq -C_1 q_n.$$

Thus

$$-C_1 w < f''^{(w)}(x) < f''^{((k+1)q_n)}(x) + C_1 q_n,$$

whence (assuming that $C_1 > 1$)

$$(6.41) \quad \left|f''^{(w)}(x)\right| \leq C_1 \max\left(w, q_n + \left|f''^{((k+1)q_n)}(x)\right|\right).$$

Let us now consider w as in the assumption. Since $f''(x+i(n, x)\alpha)$ is at least of order q_n^2 , $C' \frac{w}{q_n} f''(x+i(n, x)\alpha) \geq wq_n$ for some constant $C' > 0$. Therefore $w = O\left(\frac{w}{q_n} f''(x+i(n, x)\alpha)\right)$ and $q_n = O\left(\frac{w}{q_n} f''(x+i(n, x)\alpha)\right)$. Therefore (6.41) follows by showing

$$\left| f''^{((k+1)q_n)}(x) \right| = O\left(\frac{w}{q_n} f''(x+i(n, x)\alpha)\right).$$

But $k+1$ satisfies the assumptions of Lemma 6.8, so the above follows from (6.35) and (6.40). \square

6.7. A combinatorial lemma. We present here a combinatorial lemma which will be used in the proof of Theorem 1.3.

Lemma 6.10 (Combinatorial lemma). *Let $U, V \in \mathbb{R} \setminus \{0\}$. Fix $p, q \in \mathbb{R} \setminus \{0\}$, $p \neq q$. If $\frac{p}{q} \notin \{1, \frac{U}{V}, \frac{V}{U}\}$, then for every $K > 0$ there exists $c = c(p, q, U, V, K) > 0$ such that for every $0 < A < 10c$, for every $0 \neq B \in \mathbb{R}$ and for every $j_1, j_2 \in \{U, V\}$, we have*

$$(6.42) \quad \max(|qj_1A^{-2} - pj_2B^{-2}|, |q^2j_1A^{-3} - p^2j_2B^{-3}|) \geq K.$$

Proof. The proof goes by contradiction. Assume that for some K , there are $0 < A_n \rightarrow 0$, $0 \neq B_n \in \mathbb{R}$ and $j_{1,n}, j_{2,n} \in \{U, V\}$ such that (6.42) does not hold, that is

$$|qj_{1,n}A_n^{-2} - pj_{2,n}B_n^{-2}| < K \text{ and } |q^2j_{1,n}A_n^{-3} - p^2j_{2,n}B_n^{-3}| < K.$$

Denote $C_n = \frac{B_n}{A_n}$. Then

$$|qj_{1,n} - pj_{2,n}C_n^2| < KA_n^2 \text{ and } |q^2j_{1,n} - p^2j_{2,n}C_n^3| < KA_n^3.$$

Hence C_n is bounded, and, passing to a subsequence if necessary, we can assume that $C_n \rightarrow C$ and $j_{i,n} = j_i$ for $i = 1, 2$. It follows that $qj_1 = pj_2C^2$ and $q^2j_1 = p^2j_2C^3$. Raising to the third power the first, squaring for the second equality and dividing, we obtain $\frac{p}{q} = \frac{j_2}{j_1} \in \{1, \frac{U}{V}, \frac{V}{U}\}$, a contradiction. \square

7. DISJOINTNESS IN ARNOL'D FLOWS (PROOF OF THEOREM 1.2)

We begin by giving now an overview of the steps of the proof and the general strategy to prove Theorem 1.2.

Strategy of proof. Let $((R_\alpha)_t^f)$ be an Arnol'd flow. To prove the disjointness result in Theorem 1.2 it is enough (by Remark 2.2) to prove that for any real numbers $p, q > 0$, $p \neq q$, the special flows $((R_\alpha)_t^{pf})$ and $((R_\alpha)_t^{qf})$ are disjoint (see the beginning of Section 7.1). In order to show this, we will exploit the disjointness criterium for special flows given by Proposition 5.1 proved in Section 5.

Let us recall that Proposition 5.1 involves the definition of sets E_k (to which the initial condition x belongs), with the corresponding automorphisms A_k (so that $x' := A_k(x)$) and a set Z (to which the points y, y' belong). In Section 7.1, we first define the sets E_k so that points $x \in E_k$ do not come too close to the singularity at all (large) scales (this is made precise by (7.1)). More importantly, for $x \in E_k$, we want the q_{n_k} orbit (here q_{n_k} are the denominators of α at the special times (n_k) given by $(\mathcal{D}2)$ in the Diophantine condition, see Definition 6.1) to come close to the singularity but in a controlled way (see (7.2)), where the closeness is given by a crucial parameter $1 \gg c_{p,q} > 0$. We set $x' = A_k(x) = x + \frac{1}{q_{n_k+1}}$. Remark that, because of the definition of (n_k) , the orbit of x of length kq_n , $k \leq \frac{c_{p,q}q_{n_k+1}}{q_{n_k}}$, is almost a copy of the orbit of length q_{n_k} (we have *resonance*) and this crucially happens iterating both forward and backward.

The set Z is such that the points y, y' satisfy a Ratner-type form of shearing *either* going forward *or* backward in time (see (7.4)). This is essentially the set on which the *switchable* Ratner property (see Section 2.5) holds for the Arnold flow and the definition is indeed the same as the set Z is in [10] or [21]. If y, y' display this good form of Ratner-like shearing going *forward*, we then show that the *Forward* assumptions (F1) – (F4) in **F.** of Proposition 5.1 hold, while if the Ratner-like form of shearing happens *backward*, we show that the *Backward* assumptions (B1) – (B4) in **B.** of Proposition 5.1 hold. Notice that here it is crucial that the definition of E_k is such that x, x' have controlled shearing *both* going forward *and* backward, since the choice of whether to verify the set of assumptions **F.** or **B.** depends on the pair y, y' only.

Let us outline the proof if y, y' are *forward* 'good'. The properties (F1), (F3) and (F4) of **F.** of Proposition 5.1 all depend on a *slow* and controlled form of shearing between orbits and follow by the fact that x, x', y, y' approach the singularity in a controlled way (they do not come too close) and we may use Denjoy-Koksma type estimates (see Section 6.2). This is the content of Lemma 7.2, which is then proved in Section 7.3. The main difficulty is to prove property (F2) of Proposition 5.1, which gives the *splitting* à la Ratner of nearby trajectories. This is the content of Proposition 7.3 (in the forward case, the backward being analogous). Here the proof splits into two cases, according to what is the relative order of magnitude of the distances between the points x, x' (given by (7.7)) and between the points y, y' (see (7.6))

If **Case 1.** (called *asynchronous splitting* case) holds, these orders of magnitude are different. In this case the proof becomes simpler. We know in this case that x, x' and y, y' both display a Ratner-like shearing property going forward, but since their distances are *not* comparable, one pair will start splitting and the other one will still stay close together (depending on which distance is larger), this is why we say that the shearing is *asynchronous*. Hence property (F2) of Proposition 5.1 in this case reduces essentially to the Ratner property (forward) for one of the two flows (i.e. either $|f^{(n)}(x) - f^{(n)}(x')| \leq \epsilon$ or $|f^{(\zeta n)}(y) - f^{(\zeta n)}(y')| \leq \epsilon$ and the realignment with a shift $r_{p,q}$ is caused only by splitting for the other term).

Case 2. (dabbed *second order splitting*) corresponds to the delicate case when the distances between x, x' and y, y' are such that the shearing phenomena in the two flows, in the *main order*, are comparable and hence there is a cancellation of the main order of shearing. This is the most difficult (and technical) part of the proof. It is here that one has to understand the behaviour of the second order terms and use fine estimates on Birkhoff sums growth for further derivatives. The arguments in this case consist of two parts. The first part consists in showing that if the Birkhoff sums split by some $p \in P$ then they stay ϵ -close for a κ proportion of time (see (7.12)) and the second is that they will split at some point (see (7.13)). This part is split into two further subcases **2(a)** and **2(b)**. We give a local outline of the arguments used in this part in Section 7.4.

Let us finish the outline by giving the scale of the constants that appear: we will show that the *shifts* space is $P = \{-r_{p,q}, r_{p,q}\}$ and the constant $c_{p,q}$ involved in the definition of the sets (E_k) is such that $1 \gg c_{p,q} \gg r_{p,q} > 0$, which should be understood that $c_{p,q}$ is smaller than any algebraic expression involving constants, p and q and similarly $r_{p,q}$ is smaller than any algebraic expression involving $c_{p,q}$ (it is here that the combinatorial Lemma 6.10 is used).

7.1. Set up and beginning of the proof. Let us first remark that, for any given $\mathbb{R} \ni p, q > 0, p \neq q$, in order to show that $((R_\alpha)_p^f)^\perp \perp ((R_\alpha)_q^f)$, it is enough to prove that

$((R_\alpha)_t^{pf}) \perp ((R_\alpha)_t^{qf})$. This follows from Remark 2.2, since for a fixed $0 \neq r \in \mathbb{R}$, the flow $((R_\alpha)_{r^{-1}t}^f)$ acting on $(\mathbb{T}^f, \lambda^f)$ is isomorphic to $((R_\alpha)_t^{rf})$ acting on $(\mathbb{T}^{rf}, \lambda^{rf})$.

To show this, we will use Proposition 5.1 for f, g replaced with pf and qf , respectively. As no confusion arises, to simplify notation we will drop all ' in Proposition 5.1 (so we will have X_k, A_k, ϵ, N , etc.). From now on $p, q \in \mathbb{R}$ are fixed. We will assume that $0 < p < q$ and hence $\zeta < 1$.

To apply Proposition 5.1, we first of all need to define the sets (E_k) , the automorphisms (A_k) and the set Z which appear in it. This is the goal of this subsection. In the proof, the sets will not depend on ϵ (i.e. they are defined globally).

Definition of the sets sequences (E_k) , (X_k) and (A_k) . Let us first define the sets E_k so that each $x \in E_k$ does not come too close to the singularity at all (large) scales (this will be given by (7.1)), but the q_{n_k} orbit does come close to the singularity in a controlled way (this will be given by (7.2)). Given $s \in \mathbb{N}$, we let us first define

$$(7.1) \quad E'(s) := \left\{ x \in \mathbb{T} : x \notin \bigcup_{i=-q_s}^{q_s} R_\alpha^i \left[-\frac{1}{q_s \log^2 q_s}, \frac{1}{q_s \log^2 q_s} \right] \right\}$$

and let $E^{s_0} = \bigcap_{s \geq s_0} E'(s)$. Notice that since (q_s) grows exponentially, $\lambda(E^{s_0}) \rightarrow 1$ as $s_0 \rightarrow +\infty$.

Let (q_{n_k}) be the sequence from (D2). We now specify the important constant $c_{p,q} > 0$. We want $c_{p,q} > 0$ to be a small number, smaller than any expression involving $1, p, q$ which will appear in the course of the proof (below). In particular, we require (see Lemma 6.7 for $C'(f)$ and Lemma 6.10 for $c(\cdot, \cdot, \cdot, \cdot, \cdot)$)

$$c_{p,q} < \frac{1}{100} c(p, q, A_-, A_+, \min((C'(f) + 1)|p| + |q| + 1, (C'(f) + 1)(p^2 + q^2) + 1)).$$

Define

$$(7.2) \quad E_k := \left(\bigcup_{i=0}^{q_{n_k}^{-1}} R_\alpha^i \left[\frac{2c_{p,q}}{q_{n_k}}, \frac{3c_{p,q}}{q_{n_k}} \right] \right) \cap E^{s_0},$$

where s_0 is such that for every $k \in \mathbb{N}$, $\lambda(E_k) \geq c_{p,q}/2 > 0$.

We now define the sequences (A_k) and (X_k) . Let $\delta_k := \frac{1}{q_{n_k+1}}$, $X_k := \mathbb{T}$ and $A_k(x) = x + \delta_k$. Notice that $A_k \rightarrow Id$ uniformly and obviously $\lambda(X_k) \rightarrow \lambda(\mathbb{T})$. When k is fixed, in view of the definitions of E_k and A_k , by taking $x \in E_k$ and letting $x' = A_k x$, we obtain:

$$(7.3) \quad \left(\bigcup_{i=-[c_{p,q}q_{n_k+1}]}^{[c_{p,q}q_{n_k+1}]} R_\alpha^i [x, x'] \right) \cap \left[-\frac{c_{p,q}}{2q_{n_k}}, \frac{c_{p,q}}{2q_{n_k}} \right] = \emptyset.$$

Definition of P and Z . We first define the set $P := \{-r_{p,q}, r_{p,q}\}$, where $r_{p,q} > 0$ is chosen as follows. Let $\bar{p} = \min(p, p^{-1}, q, q^{-1}, \zeta, \zeta^{-1})$ and

$$r_{p,q} := \left(\frac{c_{p,q}\bar{p}}{100} \right)^{200}.$$

Fix $\epsilon > 0$ much smaller than $r_{p,q}$, $N \in \mathbb{N}$ and let $\kappa := \epsilon^{10}$. We now define the set $Z = Z(\epsilon, N)$ (in fact, Z will depend only on ϵ). Set

$$(7.4) \quad W_s := \left\{ y \in \mathbb{T} : y \notin \bigcup_{i=-q_s}^{q_s} R_\alpha^i \left[-\frac{1}{q_s \log^{7/8} q_s}, \frac{1}{q_s \log^{7/8} q_s} \right] \right\}$$

and let $Z := \left(\bigcap_{s \geq s_0, s \notin K_\alpha} W_s \right) \cap E^{s_0}$, where $s_0 = s_0(\epsilon)$ is such that $\lambda(Z) \geq 1 - \epsilon$ (s_0 does exist since $\lambda(W_s) \geq 1 - \frac{2}{\log^{7/8} q_s}$ and (D1) holds for α , cf. (7.1)). Let $\delta = \delta(\epsilon, N) := \frac{1}{q_{s'} \log q_{s'}}$, where $s' := [\kappa^{-20} \max\{N, s_0^2\}]$. We will moreover assume that

$$(7.5) \quad \frac{1}{\log^{1/2} \delta} < \kappa^2.$$

Assume $d(A_k, Id) < \delta$. We now take any $E_k \ni x$, set $x' := A_k x$, and let $y, y' \in Z$, $\|y - y'\| < \delta$ and we have to show that either of A. or B. in Proposition 5.1 holds for x, x', y, y' . Let $v \in \mathbb{N}$ be unique such that

$$(7.6) \quad \frac{1}{q_{v+1} \log q_{v+1}} < \|y - y'\| \leq \frac{1}{q_v \log q_v}. \quad (y\text{-variation scale})$$

By the definition of A_k and (D2), we have

$$(7.7) \quad \|x - x'\| = \frac{1}{q_{n_k+1}} \in \left[\frac{1}{q_{n_k+1} \log q_{n_k+1}}, \frac{1}{q_{n_k} \log q_{n_k}} \right]. \quad (x\text{-variation scale})$$

Let, moreover

$$(7.8) \quad T = T(x, x', y, y') := \zeta c_{p,q} \min(\|x - x'\|^{-1}, \|y - y'\|^{-1}, q_{v+1}). \quad (\text{choice of } T)$$

7.2. Forward or backward shearing and concluding arguments. We now explain the mechanism which allows one to choose whether to verify the *Forward* assumptions **F.** or the *Backward* assumptions **F.** of Proposition 5.1. We state some lemmas which correspond to the verification of these assumptions and then give the concluding arguments of the proof of Theorem 1.2.

In the two following subsections all parameters (i.e. $\epsilon, \kappa, Z, \delta$) are as defined above.

7.2.1. Forward or backward continuity. The first ingredient needed for the proof of Theorem 1.2 is the following lemma, which was proved as Lemma 4.7 in [10] and allows to show that certain orbits of nearby points, *either* in the *past*, *or* in the *future*, do not get too close to the singularity at the origin. This lemma plays a crucial role in [10] to prove the switchable Ratner (SR) property and determines whether the shearing needed in the SR-property can be seen moving *forward* (or in the future), or *backward* (or in the past).

Lemma 7.1 (Forward or backward continuity). *Fix $s \in \mathbb{N}$ and let $y, y' \in \mathbb{T}$. If at least one of the following holds:*

- (a) $s \in K_\alpha$,
- (b) $s \notin K_\alpha$ and

$$(7.9) \quad [y, y'] \cap \bigcup_{i=-q_s}^{q_s} R_\alpha^i \left[-\frac{1}{q_s \log^{7/8} q_s}, \frac{1}{q_s \log^{7/8} q_s} \right] = \emptyset,$$

then one of the following holds:

$$(7.10) \quad [y, y'] \cap \bigcup_{i=0}^{\lceil \frac{q_s+1}{4} \rceil} R_\alpha^{-i} \left[-\frac{1}{q_s \log^{7/8} q_s}, \frac{1}{q_s \log^{7/8} q_s} \right] = \emptyset.$$

or

$$(7.11) \quad [y, y'] \cap \bigcup_{i=0}^{\lceil \frac{q_s+1}{4} \rceil} R_\alpha^i \left[-\frac{1}{q_s \log^{7/8} q_s}, \frac{1}{q_s \log^{7/8} q_s} \right] = \emptyset.$$

7.2.2. Forward control. In this subsection we will assume that x, x' are as in (7.7) and satisfy (7.3), and y, y' are as in (7.6) and satisfy (7.10) (with $s = v$) and T is as in (7.8).

Remark 7.1. Notice that if $v \in K_\alpha$ then either (7.10) or (7.11) holds. We assume here that (7.10) holds, (7.11) will be treated in the next subsection. If $v \notin K_\alpha$ then the same is true provided that we show the validity of (7.9). It holds indeed as $y, y' \in Z$, cf. (7.4).

We have the following two lemmas.

Lemma 7.2. *The following is true:*

- (H1) *If t is such that $n(x, s, t) \in [\frac{T}{\log^2 T}, T] \cap \mathbb{Z}$, then property (F1) of Proposition 5.1 holds (for each $r \in \mathbb{R}$; we deal here with S^g replaced with $(R_\alpha)^{qf}$, and $n(x, s, t)$ is considered for $(R_\alpha)^{pf}$ with $s \in \mathbb{R}$ arbitrary, $(x, s) \in \mathbb{T}^{pf}$).*
- (H2) *If $w \in [0, T] \cap \mathbb{Z}$, then (F3) of Proposition 5.1 holds (here and in (H3), f, g are replaced with pf, qf , respectively).*
- (H3) *If $w, u \in [0, T] \cap \mathbb{Z}$, $|w - u| < T^{2/3}$, then (F4) of Proposition 5.1 holds.*

Let $a_w := ((pf)^{(w)}(x) - (pf)^{(w)}(x')) - ((qf)^{([\zeta w])}(y) - (qf)^{([\zeta w])}(y'))$.

Lemma 7.3. *There exists an interval $[R, S] \subset [\frac{T}{\log^2 T}, T]$ such that for all $w \in [R, S] \cap \mathbb{Z}$ such that $[w, w + \kappa w] \subset [R, S]$ and all $s \in [w, w + \kappa w] \cap \mathbb{Z}$, we have*

$$(7.12) \quad |a_w - a_s| < \epsilon^{3/2}$$

and, moreover, there exists $w_0 \in [R, (1 - \epsilon)S] \cap \mathbb{Z}$ such that

$$(7.13) \quad |a_R| < \frac{r_{p,q}}{2}, \quad |a_{w_0}| \geq r_{p,q}.$$

We will give proofs of Lemma 7.2 and Lemma 7.3 in Sections 7.3 and 7.4.

7.2.3. Backward control. In this subsection we keep unchanged our assumptions on x, x' and y, y' and T . We assume that y, y' satisfy (7.11), cf. Remark 7.1. We have the following two lemmas.

Lemma 7.4. *The following is true:*

- (H1) *If t is such that $n(x, s, -t) \in [\frac{T}{\log^2 T}, T] \cap \mathbb{Z}$, then (B1) of Proposition 5.1 holds.*
- (H2) *If $w \in [0, T] \cap \mathbb{Z}$, then (B3) of Proposition 5.1 holds.*
- (H3) *If $w, u \in [0, T] \cap \mathbb{Z}$, $|w - u| < T^{2/3}$, then (B4) of Proposition 5.1 holds.*

Let $a_w := ((pf)^{(w)}(x) - (pf)^{(w)}(x')) - ((qf)^{([\zeta w])}(y) - (qf)^{([\zeta w])}(y'))$.

Lemma 7.5. *There exists an interval $[R, S] \subset [\frac{T}{\log^2 T}, T]$ such that for all $w \in [R, S] \cap \mathbb{Z}$ such that $[w, w + \kappa w] \subset [R, S]$ and all $s \in [w, w + \kappa w] \cap \mathbb{Z}$, we have*

$$(7.14) \quad |a_{-w} - a_{-s}| < \epsilon^{3/2}$$

and, moreover, there exists $w_0 \in [R, (1 - \epsilon)S] \cap \mathbb{Z}$ such that

$$(7.15) \quad |a_{-R}| < \frac{r_{p,q}}{2}, \quad |a_{-w_0}| \geq r_{p,q}.$$

7.2.4. Concluding arguments. We now show that, using the previous lemmas, we can conclude the proof of Theorem 1.2. The next section will be then devoted only to the proof of the lemmas.

Proof of Theorem 1.2. As we have already remarked at the beginning of Section 7.1, it is sufficient to verify the assumptions of Proposition 5.1, for $(E_k), (A_k)$ and Z defined as in that section.

If (7.10) holds, we show that A. in Proposition 5.1 is true using Lemma 7.2 and Lemma 7.3. If (7.11) holds, we show that B. is true using Lemma 7.4 and Lemma 7.5. Since the proofs follow the same lines, we will assume that (7.10) is satisfied and show that A. holds.

We will show that A. in Proposition 5.1 holds. Notice that by (7.12), we have $|a_{w+1} - a_w| \leq \epsilon^{3/2}$ for $w \in [R, S]$. This and (7.13) imply that there exists $w' \in [R, (1 - \epsilon)S]$ such that $|a_{w'} \pm r_{p,q}| < 2\epsilon^{3/2}$. Define $M := w'$ and $L := \kappa w'$. Then by (7.12), for $s \in [M, M + \kappa M] \cap \mathbb{Z}$, we get

$$|a_s \pm r_{p,q}| \leq |a_s - a_{w'}| + |a_{w'} \pm r_{p,q}| \leq \epsilon^{3/2} + 2\epsilon^{3/2} < \epsilon.$$

This implies that property (F2) of Proposition 5.1 holds on $[M, M + L] \cap \mathbb{Z}$ (we recall that $P = \{r_{p,q}, -r_{p,q}\}$). Moreover, $[M, M + L] \subset [R, S] \subset [\frac{T}{\log^2 T}, T]$, whence (H1), (H2) and (H3) (in Lemma 7.2) apply in particular on $[M, M + L]$. Therefore, property (F1), (F3) and (F4) of Proposition 5.1 hold. This finishes the proof of A. \square

7.3. Slow shearing properties: proof of Lemma 7.2. In this section we give the proof of Lemma 7.2, in which the properties of the disjointness criterium which depends on *slow* and controlled shearing are verified for the *forward* case F. The proof of the parallel Lemma 7.4 for the *backward* case is analogous. In the next subsection we will prove Lemma 7.3 which verifies the crucial property (F3) (which gives the *splitting* of orbits) for the *forward* case F.

Proof of Lemma 7.2. We first show that (H2) holds. Notice that by (7.3), the x -variation scale (7.7) (indeed, $\|x - x'\|^{-1} = q_{n_k+1}$, whence $T < \zeta c_{p,q} q_{n_k+1}$), (7.10), the y -variation scale (7.6) and the choice of T in (7.8), we have

$$(7.16) \quad 0 \notin [x + w\alpha, x' + w\alpha], \quad 0 \notin [y + w\alpha, y' + w\alpha], \quad \text{for every } w \in [0, T] \cap \mathbb{Z}.$$

Therefore, for every $w \in [0, T] \cap \mathbb{Z}$, some $\theta_w \in [x, x']$, by (7.3), the x -variation scale (7.7) and (D2) together with (7.5), we have

$$|pf(R_\alpha^w x) - pf(R_\alpha^w x')| = p|f'(R_\alpha^w \theta_w)| \|x - x'\| \leq \frac{\tilde{A} p q_{n_k}}{c_{p,q}} \frac{1}{q_{n_k+1}} \leq \frac{1}{\log^{1/2} q_{n_k}} < \kappa^2,$$

where \tilde{A} is a constant depending on A_-, A_+ and g . Similarly, for some $\theta'_w \in [y, y']$, by (7.10) and the y -variation scale (7.6), we have

$$|qf(R_\alpha^w y) - qf(R_\alpha^w y')| = q|f'(R_\alpha^w \theta'_w)| \|y - y'\| \leq \tilde{A} q q_v \log^{7/8} q_v \frac{1}{q_v \log q_v} \leq$$

$$\frac{1}{\log^{1/9} q_v} < \frac{1}{\log^{1/9} q_{s'}} \leq \frac{1}{s'^{1/10}} < \kappa^2.$$

Therefore (F3) of Proposition 5.1 holds. By the cocycle identity (applied to $qf^{(u)}(y) - qf^{(w)}(y)$ and then y') and (7.16), we have that (H3) follows by showing that

$$|qf^{(w-u)}(\theta_u)| \|y - y'\| \leq \epsilon,$$

for $|w - u| \in [0, T^{2/3}] \cap \mathbb{Z}$, $u \leq T$ and some $\theta_u \in [y + u\alpha, y' + u\alpha]$. By (7.10) (since $u \leq T \leq c_{p,q}q_{v+1}$, cf. the choice of T in (7.8)), it follows that $\theta_u \notin \Sigma_v(1/4)$ (see (6.11)). Therefore, by (6.13) for $n = v$, since $T^{2/3} \leq q_{v+1}^{2/3} \leq \epsilon^4 q_v$ (cf. the assumption (D3) on the rotation number) and by the y -variation scale (7.6), we get

$$|qf^{(w-u)}(\theta_u)| \|y - y'\| \leq q\epsilon^2 q_v \log q_v \cdot \frac{1}{q_v \log q_v} = q\epsilon^2 < \epsilon.$$

This finishes the proof of (H3). It remains to show (H1). We will show the first part of property (F1) of Proposition 5.1 (i.e. the inequality involving $n(x, s, t)$), the proof of the second one following the same lines. By definition

$$(7.17) \quad -pf(x) < -s \leq t - pf^{(n(x,s,t))}(x) \leq pf(R_\alpha^{n(x,s,t)}x).$$

Let $\ell \in \mathbb{N}$ be unique such that $q_\ell \leq T < q_{\ell+1}$. We have

$$q_{\ell+1} > T > \zeta c_{p,q} \min(q_v \log q_v, q_{n_k}, q_{v+1}) > \zeta c_{p,q} \min(q_v, q_{n_k}) > \min(q_{v-\tilde{b}}, q_{n_k-\tilde{b}}),$$

where \tilde{b} is a constant (in fact, $\tilde{b} = 2 \log(\zeta c_{p,q})$), so $\ell + 1 > \min(v - \tilde{b}, n_k - \tilde{b})$. Furthermore,

$$\begin{aligned} \frac{1}{q_{s'}} &> \frac{1}{q_{s'} \log q_{s'}} = \delta > \|x - x'\| = \frac{1}{q_{n_k+1}}, \\ \frac{1}{q_{s'} \log q_{s'}} &= \delta > \|y - y'\| = \frac{1}{q_v \log q_v}, \end{aligned}$$

whence $s' < v$ and $s' < n_k + 1$. It follows that $\ell + 1 > s' - 1 - \tilde{b} > s'/2 > s_0^2$ and finally $\ell > s_0$.

Since $x \in E_k \subset E^{s_0} \subset E'(\ell + 1)$ (see (7.1) and (7.2)) and $n(x, s, t) \leq T < q_{\ell+1}$, it follows that $\min(\|x\|, \|R_\alpha^{n(x,s,t)}x\|) \geq \frac{1}{q_{\ell+1} \log^2 q_{\ell+1}}$. Therefore, by the definition of f , we have

$$\max(pf(x), pf(R_\alpha^{n(x,s,t)}x)) < \tilde{A}_1 \log q_{\ell+1} < T^{1/10}$$

cf. (D3) (of Definition 6.1). Coming back to (7.17), we have shown that

$$|t - pf^{(n(x,s,t))}(x)| < T^{1/10}.$$

Moreover, in view of (7.17), $t \geq pf^{(n(x,s,t))}(x) - T^{1/10} \geq pn(x, s, t) \inf_{\mathbb{T}} f - T^{1/10} \geq \frac{T}{\log^3 T}$, the latter inequality follows from the lower bound on $n(x, s, t)$.

Since

$$|t - n(x, s, t) \int pf d\lambda| \leq |pf^{(n(x,s,t))}(x) - n(x, s, t) \int pf d\lambda| + |t - pf^{(n(x,s,t))}(x)|$$

and the second summand is bounded by $T^{1/10}$, it is enough to show that the first summand is bounded by $T^{1/4}$ as $T^{1/4} + T^{1/10} < t^{1/3}$ (by the lower bound on t we have just shown). Therefore, we need to show that

$$|pf^{(n(x,s,t))}(x) - n(x, s, t) \int_{\mathbb{T}} pf d\lambda| < T^{1/4}.$$

This, in turn, will follow if we show that (since $\int_{\mathbb{T}} f d\lambda = 1$)

$$|f^{(n)}(x) - n| < T^{1/5},$$

for $n \in [\frac{T}{\log^2 T}, T] \cap \mathbb{Z}$. Notice that since $x \in E^{s_0} \subset E'(\ell + 1)$, for $j < T < q_{\ell+1}$, (by (D3) of Definition 6.1), we have

$$\|x + j\alpha\| \geq \frac{1}{q_{\ell+1} \log^2 q_{\ell+1}} \geq \frac{1}{q_\ell \log^2 q_\ell \cdot (2 \log^2 q_\ell)} \geq \frac{1}{2T \log^4 T}.$$

It follows that the assumptions of Lemma 6.3 are satisfied. Hence (H1) holds and the proof of Lemma 7.2 is finished. \square

7.4. Splitting of orbits: proof of Lemma 7.3. In this section we give the proof of Lemma 7.3, which gives the splitting of nearby trajectories and is the most delicate and technical part of the proof of Theorem 1.2. We give at the beginning an outline of the proof, to help the reader to go through the (sometimes complicated technically) parts of the proof.

Let us first recall that

$$a_w^p := pf^{(w)}(x) - pf^{(w)}(x'), \quad a_w^q := qf^{([\zeta w])}(y) - qf^{([\zeta w])}(y')$$

and remark that, by mean value (which can be applied thanks to (7.16)), for every $w \in [0, T] \cap \mathbb{Z}$ there exist $\theta_{w,x,x'} \in [x, x']$, $\theta_{w,y,y'} \in [y, y']$ such that

$$a_w^p = pf'^{(w)}(\theta_{w,x,x'})(x - x'), \quad \text{and} \quad a_w^q = qf'^{([\zeta w])}(\theta_{w,y,y'})(y - y').$$

The distances $\|x - x'\|$ and $\|y - y'\|$ play hence a fundamental role in comparing a_w^p and a_w^q . Recall that the denominators q_{n_k} and q_v which encode the magnitude of these distances are defined by the y -variation scale (7.6) and x -variation scale (7.7).

Outline of the proof. Let n_k and v be the indices of the denominators which encode the distances between x, x' and y, y' , respectively (defined by (7.6) and x -variation scale (7.7)). As already anticipated in the general outline given at the beginning of Section 7 that the proof of the splitting of nearby orbits has two separate cases, namely **Case 1.** (*asynchronous splitting* case) when $v = n_k$ and **Case 2.** (*second order splitting*) when $v \neq n_k$.

The asynchronous splitting case is treated first and is not so difficult, since the property (F2) of Proposition 5.1 can (in this case) be deduced by the Ratner property (forward) for one of the two flows.

The arguments in **Case 2.** consist of two parts. The first part consists in showing that if the Birkhoff sums split by some $p \in P$ then they stay ϵ -close for a κ proportion of time (see (7.12)) and the second is that they will split at some point (see (7.13)). This part is split into two further subcases **2(a)** and **2(b)**.

For this we split the cocycle inequality in property (F2) of Proposition 5.1 according to (7.27) and show (7.12) separately for $a_{1,w}$ and $a_{2,w}$ (considering two subcases **2(a)** and **2(b)**). For $a_{1,w}$ (see (7.31) and (7.33)) we use (6.15) and for $a_{2,w}$ (see (7.30), (7.32)) we use (6.14) (we have very precise estimates for the growth of the first derivative). The last part is to show (7.13) (which is also the most technical one).

In **2(a)**, we use the fact that $a_{2,w}$ dominates $a_{1,w}$ and so a_w is large for some w since by the assumptions of **I**, $a_{2,w}$ is large.

Now, in case **2(b)**, we study a_w for $w = mq_n$ (along these times we have the best control). We argue by contradiction that a_{mq_n} is always small. Then $a_{(m+l)q_n} - a_{mq_n} - a_{lq_n}$ also has to be small. We show (see (7.38) and (7.39)) that for $a_{2,\cdot}$ the above expression is small and hence we deal only with $a_{1,\cdot}$ (see (7.40)). By the cocycle identity (and estimating the third derivative), we derive (7.43) and by an analogous reasoning (using (7.43)), we get (7.44). Now, by the choice of $c_{p,q}$, it follows that the main contribution to the Birkhoff sums of the second and third derivatives is given by the closest visit. Roughly, if $\frac{A}{q_n}$ denotes the closest visit of x and $\frac{B}{q_n}$ the closest visit of y , by (7.43) and (7.44), we get that $|pA^{-2} - qB^{-2}|$

and $|p^2 A^3 \pm q^2 B^{-3}|$ both have to be small. This however cannot hold simultaneously (see Lemma 6.10). Therefore for some m , a_{mqn} is large and (7.13) holds. This concludes the outline of the proof.

In the rest of the section we now present the proof by discussing separately the two cases outlined above.

Case 1: Asynchronous shearing. We assume in this case that $n_k \neq v$. Let $\tilde{T} := \frac{T}{\log T}$. We will find $[R, S]$ which is contained in $[\frac{T}{\log^2 T}, \tilde{T}]$. By (7.10) and since $\theta_{w,y,y'} \in [y, y']$, it follows that for every $w \in [0, T] \cap \mathbb{Z}$ (see (6.11)), we have

$$(7.18) \quad \theta_{w,y,y'} \notin \Sigma_v(1/4).$$

Similarly, by (7.3) and since $\theta_{w,x,x'} \in [x, x']$, it follows that for every $w \in [0, T] \cap \mathbb{Z}$, we have

$$(7.19) \quad \theta_{w,x,x'} \notin \Sigma_{n_k}(c_{p,q}).$$

Assume first that $n_k < v$. In this case, by the choice of T in (7.8) and the y -variation scale (7.6), $T = c_{p,q} \zeta q_{n_k+1} \leq c_{p,q} q_v$. Therefore $\tilde{T} \leq \frac{q_v}{\log q_v} < \epsilon q_v$ (the function $u/\log u$ is increasing on the interval (e, ∞)) and by (6.13) together with (7.18), for $n = v$, $M = \frac{1}{4}$ and $x = \theta_{w,y,y'}$, and then by the y -variation scale (7.6), we get for $w \in [0, \tilde{T}] \cap \mathbb{Z}$

$$(7.20) \quad |a_w^q| = |q f'^{(\zeta w)}(\theta_{w,y,y'})(y - y')| \leq q \epsilon^2 q_v \log q_v \|y - y'\| \leq \epsilon^{5/3}.$$

By (6.12) in Lemma 6.4 for $n = n_k$, $M = c_{p,q}$ and $x = \theta_{w,x,x'}$, we get for $w \in [q_{n_k}, \tilde{T}] \cap \mathbb{Z}$

$$(7.21) \quad p((A_- - A_+) - \epsilon^2)w(\log w)\|x - x'\| \leq |a_w^p| \leq p((A_- - A_+) + \epsilon^2)w(\log w)\|x - x'\|.$$

Now, set

$$R := q_{n_k} \text{ and } S := \tilde{T}.$$

Take $w, s \in [R, S]$ with $s \in [w, w + \kappa w]$. We have

$$|a_w - a_s| = |(a_w^p - a_s^p) - (a_w^q - a_s^q)| \leq |a_w^p - a_s^p| + |a_w^q| + |a_s^q| \leq |a_w^p - a_s^p| + 2\epsilon^{5/3}$$

by (7.20). Now, by (7.21),

$$(7.22) \quad |a_w^p - p(A_- - A_+)w(\log w)\|x - x'\|| < p\epsilon^2 w(\log w)\|x - x'\|$$

(and the same holds for s ; if a_w^p is negative we have to modify signs). Moreover, $w < \tilde{T} = T/\log T$, whence

$$w \log w < \frac{T}{\log T} \log \left(\frac{T}{\log T} \right) < 2T = 2\zeta c_{p,q} q_{n_k+1}.$$

It follows that

$$(7.23) \quad p\epsilon^2 w(\log w)\|x - x'\| = 2p\zeta c_{p,q} \epsilon^2 < \epsilon^{5/3}$$

(remembering that $\|x - x'\| = \frac{1}{q_{n_k+1}}$). Thus

$$|a_w^p - a_s^p| = p(A_- - A_+)\|x - x'\||w \log w - s \log s| + O(\epsilon^{5/3}).$$

But $s \in [w, w + \kappa w]$, so

$$|w \log w - s \log s| \leq |w \log w - (w + \kappa w) \log(w + \kappa w)| = O(\kappa w \log w)$$

and therefore (see (7.23))

$$|a_w^p - a_s^p| = O(p(A_- - A_+)\|x - x'\|\kappa w \log w) + O(\epsilon^{5/3}) = O(\kappa) + O(\epsilon^{5/3}) < \epsilon^{3/2}$$

which completes the proof of (7.12).

Moreover, by (7.22)

$$|a_{q_{n_k}}^p - p(A_- - A_+)q_{n_k}(\log q_{n_k})\|x - x'\|| < p\epsilon^2 q_{n_k}(\log q_{n_k})\|x - x'\|.$$

But by (D3) (of Definition 6.1), $q_{n_k}(\log q_{n_k})/q_{n_k+1} \leq 1/\log \log q_{n_k}$, so $|a_R| < r_{p,q}/2$. Finally, setting $w_0 = [(1-\epsilon)\tilde{T}]$, by (7.22) and (7.23), we obtain

$$|a_{w_0}^p - p(A_- - A_+)w_0(\log w_0)\|x - x'\| < \epsilon^{5/3}.$$

But

$$\begin{aligned} w_0 \log w_0 &\asymp ((1-\epsilon)\tilde{T} \log((1-\epsilon)\tilde{T})) \asymp \\ (1-\epsilon) \frac{T}{\log T} \log \left((1-\epsilon) \frac{T}{\log T} \right) &\asymp (1-\epsilon)T \asymp q_{n_k+1}. \end{aligned}$$

Now, by (7.20) and (7.21) for a constant $D > 0$, it follows that

$$|a_{w_0}| \geq |a_{w_0}^p| - \epsilon^{5/3} \geq Dc_{p,q} - \epsilon^{5/3} > r_{p,q},$$

so (7.13) holds and the proof is finished.

Now if $n_k > v$ the proof is symmetric. One restricts the interval to $[\frac{T}{\log^2 T}, \tilde{T}] := [\frac{T}{\log^2 T}, \frac{q_{n_k}}{\log q_{n_k}}]$ then shows that (7.20) holds for a_w^f and (7.21) holds for a_w^g . This finishes the proof in case $n_k \neq v$.

Case 2: Second order splitting. $n_k = v$. Let $C_{x,y} := \min(B_{v,x}, B_{v,y})$, cf. (6.4). Notice that by (7.3) for x and (7.10) for y , we have

$$(7.24) \quad B_{v,x} \geq c_{p,q}/2 \quad \text{and} \quad B_{v,y} \geq \log^{-7/8} q_v.$$

We will find numbers R, S satisfying additionally $[R, S] \subset [\frac{T}{\log T}, \frac{C_{x,y}T}{10}]$ (notice that by (7.24) and by the choice of T in (7.8), $\frac{C_{x,y}T}{10} > 2\frac{T}{\log T}$). By (7.16), for every $w \in [0, T] \cap \mathbb{Z}$ there exist $\theta_{w,x,x'} \in [x, x']$, $\theta_{w,y,y'} \in [y, y']$ such that

$$a_w^p = pf^{(w)}(x)(x - x') + \frac{p}{2}f''^{(w)}(\theta_{w,x,x'})(x - x')^2$$

and

$$a_w^q = f'^{([\zeta w])}(y)(y - y') + \frac{q}{2}f''^{([\zeta w])}(\theta_{w,y,y'})(y - y')^2.$$

Notice that since $\theta_{w,x,x'} \in [x, x']$, by (7.3) and (7.7), we have $B_{v,\theta_{w,x,x'}} \geq \frac{c_{p,q}}{2}$ for every $w \in [0, \frac{C_{x,y}T}{10}] \cap \mathbb{Z}$. So, by (6.39) in Lemma 6.9, for every $w \in [0, \frac{C_{x,y}T}{10}] \cap \mathbb{Z}$ (recall the choice of T in (7.8))

$$(7.25) \quad |f''^{(w)}(\theta_{w,x,x'})|(x - x')^2 \leq DC_{x,y}Tq_v \frac{4}{c_{p,q}^2} \frac{1}{q_{v+1}} \leq \epsilon^{10}$$

(here, we use the fact that $v = n_k$ satisfies (D2)). Similarly, since $\theta_{w,y,y'} \in [y, y']$, by (7.10) and the y -variation scale (7.6), we have $B_{v,\theta_{w,y,y'}} \geq \frac{1}{\log^{7/8} q_v}$ for every $w \in [0, \frac{C_{x,y}T}{10}] \cap \mathbb{Z}$. So by (6.39) in Lemma 6.9, for every $w \in [0, \frac{C_{x,y}T}{10}] \cap \mathbb{Z}$, since $C_{x,y} \leq B_{v,y}$ and $B_{v,y} \geq \log^{-7/8} q_v$, we have (recall the choice of T in (7.8))

$$(7.26) \quad f''^{([\zeta w])}(\theta_{w,y,y'})(y - y')^2 \leq \frac{\zeta 4DC_{x,y}Tq_v B_{v,y}^{-2}}{Tq_v \log q_v} = O\left(\frac{\log^{7/8} q_v}{\log q_v}\right) \leq \epsilon^{10}.$$

Since $a_w = a_w^p + a_w^q$, by (7.25), (7.26), it is enough to study the expression $\bar{a}_w = pf^{(w)}(x)(x - x') - qf'^{([\zeta w])}(y)(y - y')$ for $w \in [0, \frac{C_{x,y}T}{10}]$. We rewrite it in the form

$$(7.27) \quad \bar{a}_w = a_{1,w} + a_{2,w}, \quad \text{where} \quad a_{1,w} := \left(pf^{(w)}(x) - qf'^{([\zeta w])}(y)\right)(y - y'), \\ a_{2,w} := pf^{(w)}(x)((x - x') - (y - y')).$$

We consider now two subcases (Case 2(a) and Case 2(b)):

Case 2(a). In this subcase, we assume that there exists $w_0 \in [\frac{T}{\log T}, \frac{C_{x,y}T}{30}] \cap \mathbb{Z}$ such that

$$|a_{2,w_0}| \geq 10.$$

Notice that by the choice of T in (7.8), $\frac{T}{\log T} \geq c_{p,q}\zeta q_v$. Therefore and since $x \notin \Sigma_v(c_{p,q}/2)$ (cf. (7.3)) by (6.12), (7.7) and the y -variation scale (7.6), we have

$$|a_{2, [\frac{T}{\log T}]}| \leq 2p|A_- - A_+|T^{\frac{2c_{p,q}}{T}} \leq 5.$$

Moreover, for every $w \in [\frac{T}{\log T}, \frac{C_{x,y}T}{10}] \cap \mathbb{Z}$ by (7.3), (7.7) and the y -variation scale (7.6), we have

$$|a_{2,w+1} - a_{2,w}| = p|f'(x + w\alpha)| |(x - x') - (y - y')| \leq 4|A_- - A_+|pc_{p,q}^{-1}q_v \frac{2}{q_v \log q_v} < \epsilon^3$$

and therefore there exists $w' \in [\frac{T}{\log T}, \frac{C_{x,y}T}{30}] \cap \mathbb{Z}$ such that

$$(7.28) \quad |a_{2,w'} \pm 10| < \epsilon^2.$$

In particular, using additionally (6.12), we have

$$(7.29) \quad w' \log w' \|(x - x') - (y - y')\| = O(1).$$

Define $R = w'$, $S := (1 + \epsilon)w'$. Then for every $w \in [R, S]$ such that $[w, w + \kappa w] \subset [R, S]$ and every $s \in [w, w + \kappa w]$, by (6.12) and (7.29), we have

$$(7.30) \quad |a_{2,w} - a_{2,s}| \leq (p|A_- - A_+|(s \log s - w \log w) + O(\epsilon^2 w \log w + \epsilon^2 s \log s)) \|(x - x') - (y - y')\| = O(\kappa w' \log w') \|(x - x') - (y - y')\| + O(\epsilon^2) = O(\epsilon^2).$$

Moreover, by (6.15) (for x and y), we have (recall that $\zeta = p/q$)

$$(7.31) \quad |a_{1,w} - a_{1,s}| = \|y - y'\| \left| (pf'^{(w)}(x) - pf'^{(s)}(x)) - (qf'^{([\zeta w])}(y) - qf'^{([\zeta s])}(y)) \right| = O\left(\frac{\epsilon^2 T}{T}\right) = O(\epsilon^2).$$

By the two above, for every $w \in [R, S]$, $s \in [w, w + \epsilon^3 w]$, we have

$$|a_w - a_s| \leq \epsilon^{3/2}.$$

which gives (7.12). Moreover, by (6.14) (recalling that $\zeta = \frac{p}{q}$) and by the choice of T in (7.8), we have for w'

$$|a_{w'}| \geq |a_{2,w'}| - |a_{1,w'}| \geq 10 - \epsilon^2 - \|y - y'\|T \geq 8 > r_{p,q},$$

and so (7.13) holds. This finishes the proof in case **1(a)**.

Case 2(b). In this case, we assume that, contrary to **Case 2(a)**, for every $w \in [\frac{T}{\log T}, \frac{C_{x,y}T}{30}] \cap \mathbb{Z}$, we have

$$|a_{2,w}| \leq 10.$$

This, by (6.12) (since $x \notin \Sigma_l$ by (7.3)), means that for every $w, s \in [\frac{T}{\log T}, \frac{C_{x,y}T}{30}] \cap \mathbb{Z}$, $s \in [w, w + \epsilon^3 w]$, we have (similarly to **1(a)**)

$$(7.32) \quad |a_{2,w} - a_{2,s}| \leq \epsilon^{5/3}.$$

Moreover, for $w, s \in [\frac{T}{\log T}, \frac{C_{x,y}T}{30}]$, $s \in [w, w + \epsilon^3 w]$, by (6.15) and the definition of $a_{1,w}$, we have (similarly to $\mathbf{1}(\mathbf{a})$.)

$$(7.33) \quad |a_{1,w} - a_{1,s}| \leq \epsilon^{5/3}.$$

So, by (7.32) and (7.33), it follows that (7.12) holds for every interval $[R, S] \subset [\frac{T}{\log T}, \frac{C_{x,y}T}{30}] \cap \mathbb{Z}$. Therefore to finish the proof it is enough to show that there exists $w_0 \in [\frac{T}{\log T}, \frac{C_{x,y}T}{40}]$ such that (7.13) holds.

By assumption, $10 \geq |a_{2,w}| = |pf'^{(w)}(x)| |(x - x') - (y - y')|$. Moreover, by the choice of T in (7.8), $T \geq c_{p,q} \zeta q_v \log q_v$. Therefore, by (6.12) (notice that $x \notin \Sigma_l$ by (7.3)) for $w = [\frac{C_{x,y}T}{40}]$, we get (using (7.24) and (D3).)

$$(7.34) \quad 10 \geq \frac{p}{100} |(x - x') - (y - y')| C_{x,y} T \log T \geq \frac{pc_{p,q}\zeta}{200} q_v \log^2 q_v |(x - x') - (y - y')| \geq q_{v+1} \log(\log q_v) |(x - x') - (y - y')|.$$

This, by the x -scale (7.7) and the y -scale (7.6), means that

$$(7.35) \quad \frac{1}{2q_{v+1}} < \|y - y'\| < \frac{2}{q_{v+1}}.$$

So, by the choice of T in (7.8), we get

$$(7.36) \quad \frac{\zeta c_{p,q} q_{v+1}}{2} < T < 2\zeta c_{p,q} q_{v+1}.$$

We will show that (7.13) holds for $w_0 = qmq_v$ for some $m \in J_{v,T} := [\frac{T}{qq_v \log T}, \frac{C_{x,y}T}{40qq_v}] \cap \mathbb{Z}$. This follows by showing that for some $m, l \in J_{v,T}$ such that $m + l \in J_{v,T}$, we have

$$(7.37) \quad |a_{q(m+l)q_v} - a_{qmq_v} - a_{qlq_v}| > 3r_{p,q}.$$

Notice that, by definition of $a_{2,w}$, we have for some $\theta' \in [x, x + qmq_v\alpha]$ (by the cocycle identity)

$$(7.38) \quad |a_{2,q(m+l)q_v} - a_{2,qmq_v} - a_{2,qlq_v}| = |pf''^{(qlq_v)}(\theta')qm\|q_v\alpha\|((x - x') - (y - y'))|$$

Notice that $m \in J_{v,T}$, so, by (7.36), it follows that $qm\|q_v\alpha\| \leq \frac{4c_{p,q}^2}{2}$ and therefore since $\theta' \in [x, x + qmq_v\alpha]$ and (7.3) holds, we have $B_{u,\theta'} \geq \frac{c_{p,q}}{2}$. Therefore, by Lemma 6.8 and (7.34), we obtain

$$(7.39) \quad p|f''^{(qlq_v)}(\theta')|qm\|q_v\alpha\|((x - x') - (y - y'))| \leq \frac{80pq^2mlC(f)c_{p,q}^{-2}q_v^2}{q_{v+1}^2 \log(\log q_v)} \leq \epsilon,$$

where the last inequality holds since $m, l \in J_{v,T}$ so in particular (by (7.36)) $m, l \leq \frac{q_v+1}{40qq_v}$. Hence (7.37) follows by showing

$$(7.40) \quad |a_{1,q(m+l)q_v} - a_{1,qmq_v} - a_{1,qlq_v}| > 4r_{p,q}.$$

We argue by contradiction assuming that for every $m, l \in J_{v,T}$ such that $m + l \in J_{v,T}$, (7.40) does not hold.

We will often use the following condition: If $\theta \in [x, x + qmq_v\alpha]$, $\theta' \in [y, y + \zeta pmq_v\alpha]$ and $m \in J_{v,T}$, then

$$(7.41) \quad 2B(v, x) \geq B_{v,\theta} \geq \frac{B_{v,x}}{2} \text{ and } 2B(v, y) \geq B_{v,\theta'} \geq \frac{B_{v,y}}{2}.$$

We will first show that

$$(7.42) \quad C(f, p, q)c_{p,q} \leq C(f, p, q)B_{v,x} < B_{v,y} < C(f, p, q)^{-1}B_{v,x}.$$

for some constant $C(f, p, q) > 0$, with $C(f, p, q)^9 \geq r_{p,q}$. Notice that by (7.24), the right inequality is immediate ($B_{v,y} \leq 1$ and $B_{v,x} \geq c_{p,q}$). If $B_{v,y} \geq c(f)$ (see Lemma 6.8) then (7.42) follows (since $B(v, x) < 1$). Hence, we will assume that $B(v, y) \leq c(f)$ so that we can use Lemma 6.8 for y and v . Notice that we can use Lemma 6.8 for x since $B_{v,x} \leq c_{p,q} \leq c(f)$ (see (7.2)). Notice also that by the cocycle identity, for some $\theta_{m,l} \in [x, x + qmq_v\alpha]$,

$$pf^{(q(m+l)q_v)}(x) - pf^{(qmq_v)}(x) - pf^{(qlq_v)}(x) = pf^{''qlq_v}(\theta_{m,l})qm\|q_v\alpha\|.$$

and for some $\theta'_{m,l} \in [y, y + \zeta qmq_v\alpha]$, we get (using that $\zeta = p/q$)

$$qf^{(q(m+l)q_v)}(x) - qf^{(qmq_v)}(x) - qf^{(qlq_v)}(x) = qf^{''qlq_v}(\theta'_{m,l})pm\|q_v\alpha\|.$$

Hence, by the definition of $a_{1,w}$ (recall that, for sake of contradiction, we assume that (7.40) does not hold), it follows that

$$4r_{p,q} > \|y - y'\| \left| pf^{''qlq_v}(\theta_{m,l})qm\|q_v\alpha\| - qf^{''plq_v}(\theta'_{m,l})pm\|q_v\alpha\| \right|.$$

This, for $m = l = \lceil \frac{C_{x,y}T}{80q q_v} \rceil$ (then $m, l, m + l \in J_{v,T}$), imply by (7.35), (7.41) (for $\theta_{m,l}$ and $\theta'_{m,l}$), (6.35) and (7.36) that

$$\begin{aligned} \frac{16r_{p,q}}{pq} q_{v+1}^2 &\geq m |f^{''qlq_v}(\theta_{m,l}) - f^{''plq_v}(\theta'_{m,l})| \geq ml q_v^2 (C(f)pB_{v,y}^{-2} - c(f)qB_{v,x}^{-2}) \geq \\ & q_{v+1}^2 \frac{\zeta^2 c_{p,q}^2}{10^6 q} C_{x,y}^2 (C(f)pB_{v,y}^{-2} - c(f)qB_{v,x}^{-2}). \end{aligned}$$

This, by the definition of $C_{x,y}$, finishes the proof of (7.42) since if it does not hold then $C_{x,y} = B_{v,y}$ and then the inequality above shows that

$$C(f)p \leq 3c(f)qc_{p,q}^{-3}B_{v,y}^2 + \frac{10^8 qr_{p,q}}{\zeta^2 c_{p,q}^2},$$

which is not true since $r_{p,q} > 0$ is taken sufficiently small. Hence (see (7.24)), $C_{x,y} \geq c_{p,q} \min(1, C(f, p, q))$ and so the right endpoint of $J_{v,T}$ is a fixed (depending on p, q only) proportion of q_{v+1} (see (7.36)). To emphasize that denote $C_{x,y} := C_{p,q}$.

Notice that for some $\theta_{m,l} \in [x, x + qmq_v\alpha]$, we have

$$pf^{(q(m+l)q_v)}(x) - pf^{(qmq_v)}(x) - pf^{(qlq_v)}(x) = pf^{''qlq_v}(x)qm\|q_v\alpha\| + pf^{'''qlq_v}(\theta_{m,l})(qm\|q_v\alpha\|)^2.$$

Moreover, by (7.41) and (6.36), we get

$$pf^{'''qlq_v}(\theta_{m,l})(qm\|q_v\alpha\|)^2 \leq \frac{C'(f)pq^3m^2lq_v^3C_{p,q}^{-3}}{q_{v+1}^2}$$

Similarly, for some $\theta'_{m,l} \in [y, y + \zeta qmq_v\alpha]$, we have

$$qf^{(p(m+l)q_v)}(y) - qf^{(pmq_v)}(x) - qf^{(plq_v)}(x) = qf^{''plq_v}(x)pm\|q_v\alpha\| + qf^{'''plq_v}(\theta'_{m,l})(pm\|q_v\alpha\|)^2$$

and by (7.41) and (6.36), we get

$$qf^{'''plq_v}(\theta'_{m,l})(pm\|q_v\alpha\|)^2 \leq \frac{C'(f)qp^3m^2lq_v^3C_{p,q}^{-3}}{q_{v+1}^2}.$$

Therefore, by the definition of $a_{1,w}$ and by (7.40), for all $m, l, m + l \in J_{v,T}$, we have (since $p < q$ and so $p^3q < pq^3$)

$$(7.43) \quad |y - y'|m\|q_v\alpha\|pq|f^{''qlq_v}(x) - f^{''plq_v}(y)| \leq 10r_{p,q} + 2 \frac{C'(f)pq^3m^2lq_v^3C_{p,q}^{-3}}{q_{v+1}^3}.$$

Denote the RHS of (7.43) by $e_1(l, m)$.

Let $h, l \in J_{v,T}$ be such that $m + l + h \in J_{v,T}$. Notice that for $b \in \{p, q\}$, for some $\theta_{h,l}^b \in [x, x + qlq_v\alpha]$, we have

$$f''^{(b(l+h)q_v)}(x) - f''^{(blq_v)}(x) - f''^{(bhq_v)}(x) = f'''^{(bhq_v)}(x)bl\|q_v\alpha\| + f''''^{(bhq_v)}(\theta_{h,l}^b)(bl\|q_v\alpha\|)^2.$$

By (7.41) for $\theta_{h,l}^b$ and (6.37), we have

$$f''''^{(bhq_v)}(\theta_{h,l}^b)(bl\|q_v\alpha\|)^2 \leq \frac{4C'(f)b^3l^2hq_v^4C_{p,q}^{-4}}{q_{v+1}^2}.$$

Using this for $b = p$ and $b = q$, in view of (7.43), for $l + h, l, h$, we have

$$(7.44) \quad |y - y'|ml\|q_v\alpha\|^2|qf'''^{(qhqv)}(x) - pf'''^{(phqv)}(x)| \leq e_2(m, l, h),$$

where

$$|e_2(m, l, h)| \leq 6 \left(10r_{p,q} + 2 \frac{C'(f)pq^3m^2lq_v^3C_{p,q}^{-3}}{q_{v+1}^3} \right) + \frac{20C'(f)pq(p^4 + q^4)hml^2q_v^4C_{p,q}^{-4}}{q_{v+1}^4}.$$

We will show that (7.43) and (7.44) can not both be true. Consider l, m, h for which $\max(p, q, 1) \max(l, m, h) \leq C_{p,q}^6$. Then, by (7.42), the assumptions of Lemma 6.7 are satisfied for x, y, v . So, by (6.32), for some $j(x, v), j(y, v) \in \{A_-, A_+\}$, we have

$$|f''^{qlqv}(x) - f''^{plqv}(y)| \geq |qlq_v^2j(x, v)B_{v,x}^{-2} - plq_v^2j(y, v)B_{v,y}^{-2}| - 2(C'(f) + 1)(|p| + |q|)lq_v^2$$

and by (6.33)

$$\begin{aligned} & |qf'''^{qhqv}(x) - pf'''^{phqv}(x)| \geq \\ & |2j(x, v)q^2hq_v^3B_{v,x}^{-3} \pm 2j(y, v)p^2hq_v^3B_{v,y}^{-3}| - 2(p^2 + q^2 + 1)(C'(f) + 1)qhqv^3. \end{aligned}$$

Therefore in (7.43), using (7.35), we have

$$(7.45) \quad |qj(x, v)B_{v,x}^{-2} - pj(y, v)B_{v,y}^{-2}| - 2(C(f) + 1)(|p| + |q|) \leq e_1(m, l) \frac{q_{v+1}^2}{pqmlq_v^2}$$

and

$$(7.46) \quad |2q^2j(x, v)B_{v,x}^{-3} \pm p^2j(y, v)B_{v,y}^{-3}| - 2(C'(f) + 1)(p^2 + q^2) \leq e_2(m, l, h) \frac{q_{v+1}^3}{pqmlhq_v^3}.$$

Let $m = \lfloor \frac{Jq_{v+1}}{q_v} \rfloor, l = \lfloor \frac{Kq_{v+1}}{q_v} \rfloor, h = \lfloor \frac{Wq_{v+1}}{q_v} \rfloor$. Then, by the definition of $e_1(m, l)$, we have

$$|e_1(m, l) \frac{q_{v+1}^2}{pqmlq_v^2}| \leq \frac{r_{pq}}{pqJK} + 2JC'(f)q^2C_{p,q}^{-3}$$

and

$$|e_2(m, l, h) \frac{q_{v+1}^3}{pqmlhq_v^3}| \leq 6 \left(\frac{10r_{p,q}}{pqJKW} + \frac{2C'(f)q^2JC_{p,q}^{-3}}{W} \right) + 20C'(f)(p^4 + q^4)KC_{p,q}^{-4}.$$

If we now choose $J := \lfloor \frac{C_{p,q}^6}{2C'(f)q^2} \rfloor, K := \lfloor \frac{C_{p,q}^4}{20C'(f)(p^4 + q^4)} \rfloor$ and $W := \lfloor C_{p,q}^3 \rfloor$ and remember that $r_{p,q} < (pq)^{-1}C_{p,q}^{30}$ then, by the above estimates and (7.45), (7.46), we have

$$|qj(x, v)B_{v,x}^{-2} - pj(y, v)B_{v,y}^{-2}| \leq 2(C(f) + 1)(|p| + |q|) + 10$$

and

$$|q^2j(x, v)B_{v,x}^{-3} \pm p^2j(y, v)B_{v,y}^{-3}| < 2(C(f) + 1)(p^2 + q^2) + 10.$$

If in the latter inequality “ \pm ” is in fact “ $+$ ”, then

$$\begin{aligned} & |q^2j(x, v)B_{v,x}^{-3} + p^2j(y, v)B_{v,y}^{-3}| \geq q^2j(x, v)B_{v,x}^{-3} \geq \\ & 10^{-3}q_2 \min(A_-, A_+)c_{p,q}^{-3} > 2(C(f) + 1)(p^2 + q^2) + 10, \end{aligned}$$

by the choice of $c_{p,q} > 0$ and this gives a contradiction. So we have

$$|qj(x, v)B_{v,x}^{-2} - pj(y, v)B_{v,y}^{-2}| \leq 2(C(f) + 1)(|p| + |q|) + 10$$

and

$$|q^2 j(x, v)B_{v,x}^{-3} - p^2 j(y, v)B_{v,y}^{-3}| < 2(C(f) + 1)(p^2 + q^2) + 10.$$

But, by Lemma 6.10 for $U = A_-$ and $V = A_+$, this means (since $B_{v,x} \leq c_{p,q}$) that $\frac{p}{q} \in \{1, \frac{A_-}{A_+}, \frac{A_+}{A_-}\}$ which is a contradiction with the assumptions on p, q . This contradiction shows that (7.40) and hence also (7.37) holds, which in turn implies (7.13). This finishes the proof.

8. DISJOINTNESS OF TIME CHANGES OF HOROCYCLE FLOWS AND ARNOL'D FLOWS (PROOF OF THEOREM 1.3)

In this section we will prove Theorem 1.3. Since $\tau \in C^1(M)$ is fixed, we denote (\tilde{h}_t^τ) by (\tilde{h}_t) . We will divide the proof into several steps. We will start with the following general lemma. We recall that d^f stands for the product metric.

Lemma 8.1. *For every $0 < \epsilon < 1/100$, if $d^f((y, s), (y', s')) < \epsilon^3$ and $t \in \mathbb{R}$ is such that $((R_\alpha)^f)_t(y, s) \in \{(x, u) \in \mathbb{T}^f : \epsilon^2 < u < f(x) - \epsilon^2\}$ and for $w \in \{n(y, s, t), n(y, s, t) + 1\}$, we have*

$$|f^{(w)}(y) - f^{(w)}(y') - \tilde{b}(t)| < 4\epsilon^{8/3}$$

for some function $\tilde{b} : \mathbb{R} \rightarrow \mathbb{R}$, then

$$d^f \left((R_\alpha)_t^f(y, s), (R_\alpha)_{t-\tilde{b}(t)}^f(y', s') \right) < \epsilon^2.$$

Proof. We will show the proof in case $t > 0$; the proof of case $t < 0$ follows the same lines. By assumptions

$$f^{(n(y,s,t))}(y') \leq f^{(n(y,s,t))}(y) - \tilde{b}(t) + 4\epsilon^{8/3} \leq t - \epsilon^2 + s' - \tilde{b}(t) + |s - s'| + 4\epsilon^{8/3} \leq t - \tilde{b}(t) + s'$$

and

$$f^{(n(y,s,t)+1)}(y') \geq f^{(n(y,s,t)+1)}(y) - \tilde{b}(t) - 4\epsilon^{8/3} \geq t + \epsilon^2 + s' - \tilde{b}(t) - |s - s'| - 4\epsilon^{8/3} \geq t + s' - \tilde{b}(t).$$

Therefore, $(R_\alpha)_{t-\tilde{b}(t)}^f(y', s') = (R_\alpha^{n(y,s,t)})_{t-\tilde{b}(t)}^f(y', s')$ and, by definition, $(R_\alpha)_t^f(y, s) = (R_\alpha^{n(y,s,t)})_t^f(y, s)$. This, by our assumptions, finishes the proof. \square

The set \mathcal{D}' . Let us define the set \mathcal{D}' by

$$\mathcal{D}' := \{\alpha \in [0, 1] \setminus \mathbb{Q} : (\mathcal{D}1) \text{ and } (\mathcal{D}3) \text{ (of Definition 6.1) are satisfied}\}.$$

For the proof of Theorem 1.3, we will show that the assumptions of Theorem 3.1 are satisfied for $(T_t) = (\tilde{h}_t)$ and $(S_t) = ((R_\alpha)_t^f)$. To simplify notation, we assume that $\int_{\mathbb{T}} f d\lambda = 1$.

The set P . Let $\bar{A} = \max(|A_- - A_+|^{\pm 1})$ and define

$$P := \left\{ \frac{1}{32 \cdot 2018 \bar{A}^2}, \frac{1}{32 \cdot 2018 \bar{A}^2} \right\}.$$

Construction of (X_k) and (A_k) . For $k \in \mathbb{N}$, set $X_k := M$ and $A_k x := v_{1/k} x$ (here (v_t) denotes the opposite horocycle flow). Obviously $A_k \rightarrow Id$ uniformly on M .

Construction of (E_k) . In the statement of Theorem 3.1 the sets $(E_k(\epsilon))$ depend on the parameter $\epsilon > 0$. However in our case we simply set $E_k(\epsilon) = M$ for every $k \in \mathbb{N}$ and $\epsilon > 0$.

Fix $\epsilon > 0$ and $N \in \mathbb{N}$. Let $\kappa = \kappa(\epsilon) = \epsilon^{20}$. By ergodic theorem it follows that there exist $V_\epsilon > 0$ and a set $\overline{W}_\epsilon \subset \mathbb{T}^f$, $\lambda^f(\overline{W}_\epsilon) > 1 - \epsilon^2$ such that for every $(y, s) \in \overline{W}_\epsilon$ and every $|V| > V_\epsilon$, $U \geq \kappa|V|$, we have

$$(8.1) \quad \lambda \left(\left\{ t \in [V, V+U] : ((R_\alpha)^f)_t(y, s) \in \{(x, s) \in \mathbb{T}^f : \epsilon^2 < s < f(x) - \epsilon^2\} \right\} \right) \geq (1 - \epsilon^{5/3})U$$

and for every $(y, s) \in W_\epsilon$ and $|t| \geq V_\epsilon$, we have

$$(8.2) \quad |n(y, s, t) - t| < \epsilon^3|t|.$$

We will now define a set $Z_\epsilon \subset \mathbb{T}$. First, let

$$W_s := \left\{ y \in \mathbb{T} : y \notin \bigcup_{i=-q_s}^{q_s} R_\alpha^i \left[-\frac{1}{q_s \log^{7/8} q_s}, \frac{1}{q_s \log^{7/8} q_s} \right] \right\}.$$

and set $Z(u) := \bigcap_{s \geq u, s \notin K_\alpha} W_s$. Since $\lambda(W_s) \geq 1 - \frac{2}{\log^{7/8} q_s}$, by (D1), it follows that $\lim_{u \rightarrow +\infty} \lambda(Z(u)) = 1$. Let $u_\epsilon \in \mathbb{N}$ be such that $Z(u_\epsilon) > 1 - \epsilon^2$ and define $Z_\epsilon := Z(u_\epsilon)$. We then set

$$Z := \overline{W}_\epsilon \cap Z_\epsilon^f \subset \mathbb{T}^f.$$

By the definition of f , we have $\lambda^f(Z_\epsilon^f) \geq 1 - \epsilon^{5/3}$. Therefore, $\lambda^f(Z) \geq 1 - \epsilon$.

Finally, let $\delta = \delta(\epsilon, N) := \min \left(\kappa^{20}, V_\epsilon^{-4}, \frac{1}{q_{s'} \log q_{s'}}, \bar{\delta}(\epsilon^2), N_{\epsilon^3}^{-3}, n_0(\epsilon^2)^2 \right)$, where $s' := \kappa^{-1} \max(N, u_\epsilon^2)$ and $\bar{\delta}(\epsilon^2), N_{\epsilon^3}^{-3}$ come from Proposition 4.1 (with $K = 1$) and $n_0(\epsilon^2)$ comes from Lemma 6.4. We now take $x, x' := A_k x$ with $d_M(A_k, Id) \leq \delta$ and $(y, s), (y', s') \in Z$ with $d^f((y, s), (y', s')) < \delta$ and we will show that (3.2) or (3.3) holds for x, x' and $(y, s), (y', s')$.

Let $v \in \mathbb{N}$ be unique such that

$$(8.3) \quad \frac{1}{q_{v+1} \log q_{v+1}} \leq \|y - y'\| \leq \frac{1}{q_v \log q_v},$$

and let $\ell \in \mathbb{R}$, $\ell \leq q_{v+1}$ be unique such that

$$(8.4) \quad \|y - y'\| = \frac{1}{\ell \log \ell}.$$

By the definition of Z (in particular by the definition of W_v), it follows that the assumption of Lemma 7.1 are satisfied for $s = v$ and $y, y' \in \mathbb{T}$. If (7.10) holds, we will show (3.2) and if (7.11) holds we will show (3.3). Since the proofs of both cases are analogous, we will assume that (7.10) is true for y, y' . Let

$$(8.5) \quad \tilde{T} := \frac{1}{8} \min(k^{1/2}, \ell).$$

By Proposition 4.1 (for $K = 1$ and ϵ^2) and (8.5), it follows that for every $t \in [N_\epsilon, \tilde{T}]$, we have that (4.5) holds for x, x' (with ϵ^2 instead ϵ). Since $x' = v_{k^{-1}t}x$, by the definition of $\chi(\cdot) = \chi_{x, x'}(\cdot)$, it follows that $\chi(t) = t - k^{-1}t^2$ (cf. (4.2)) and moreover, by (4.4), we get that $|A_x(t)| \leq \epsilon^2$ (since $s = 0$ for x and x') for $t \in [N_\epsilon, \tilde{T}]$. Therefore, by (4.5), we have

$$(8.6) \quad d_M(\tilde{h}_t x, \tilde{h}_{t-k^{-1}t^2} x') < 2\epsilon^2 \quad \text{for } t \in [N_\epsilon, \tilde{T}].$$

Notice that by (7.10) and (8.5) (recall that $\ell \leq q_{v+1}$) it follows that for every $n \in [0, \tilde{T}]$, by the mean value theorem, we have

$$(8.7) \quad f^{(n)}(y) - f^{(n)}(y') = f'^{(n)}(\theta_n)(y - y'),$$

where $\theta_n \in [y, y']$. Moreover, since $\theta_n \in [y, y']$, it follows by (7.10) that $\theta_n \notin \Sigma_v(1/4)$ (see (6.11)) and hence we can use Lemma 6.4. Therefore, for every $n \in [\epsilon^4 q_v, \tilde{T}]$, by (8.7), (8.4) and (6.12), we have

$$(8.8) \quad \left| f^{(n)}(y) - f^{(n)}(y') - (A_- - A_+) \frac{n \log n}{\ell \log \ell} \right| = O(\epsilon^2).$$

Similarly, by (8.7), (8.4) and (6.13) for every $n \in [0, \epsilon^4 q_v]$, we have

$$(8.9) \quad |f^{(n)}(y) - f^{(n)}(y')| = O(\epsilon^2).$$

Notice that by (8.3) and (8.4), for $n \in [0, \epsilon^4 q_v]$, we have $|(A_- - A_+) \frac{n \log n}{\ell \log \ell}| = O(\epsilon^4)$. Therefore and by (8.8) and (8.9) for every $n \in [0, \tilde{T}]$, we have

$$(8.10) \quad \left| f^{(n)}(y) - f^{(n)}(y') - (A_- - A_+) \frac{n \log n}{\ell \log \ell} \right| = O(\epsilon^2).$$

Moreover, since $(y, s) \in Z^f$, by (8.2), for $t \leq (1 - \epsilon^2)\tilde{T}$, we have

$$n(y, s, t) \leq \tilde{T} - 1$$

and using (8.2) again, we have

$$|(A_- - A_+) \frac{n(y, s, t) \log n(y, s, t)}{\ell \log \ell} - (A_- - A_+) \frac{t \log t}{\ell \log \ell}| \leq \frac{|A_- - A_+|}{\ell \log \ell} 2\epsilon^3 t \log t = O(\epsilon^3),$$

the last inequality by (8.5). The two above inequalities and (8.10) imply that for every $t \in [0, (1 - \epsilon^2)\tilde{T}]$, for $w_t \in \{n(y, s, t), n(y, s, t) + 1\}$, we have

$$(8.11) \quad |f^{(w_t)}(y) - f^{(w_t)}(y') - \tilde{b}(t)| \leq 2\epsilon^{8/3},$$

where $\tilde{b}(t) = (A_- - A_+) \frac{t \log t}{\ell \log \ell}^{10}$.

But Lemma 8.1 (cf. (8.11)) implies that for every $t \in [0, \tilde{T}]$ for which $(R_\alpha)_t^f(y, s) \in \{(y, s) \in \mathbb{T}^f : \epsilon^2 < s < f(y) - \epsilon^2\}$, we have (since $d^f((y, s), (y', s')) < \delta < \epsilon^3$)

$$(8.12) \quad d^f((R_\alpha)_t^f(y, s), (R_\alpha)_{t-\tilde{b}(t)}^f(y', s')) < \epsilon^2.$$

Let $\tilde{c}(t) := -k^{-1}t^2 + \tilde{b}(t)$. We claim the following:

$$(8.13) \quad |\tilde{c}(T_0)| > \frac{1}{32 \cdot 2018 \bar{A}^2}, \quad \text{for some } T_0 \in \left[\frac{\tilde{T}}{8\bar{A}}, \frac{\tilde{T}}{2\bar{A}} \right].$$

Indeed, denote $K_0 := \frac{\tilde{T}}{8\bar{A}}$. Then, notice that by (8.5), we have

$$|\tilde{c}(4K_0) - 16\tilde{c}(K_0) + 12(A_- - A_+) \frac{K_0 \log K_0}{\ell \log \ell}| = |A_- - A_+| \frac{4K_0 \log 4K_0 - 4K_0 \log K_0}{\ell \log \ell} < \epsilon^2,$$

and

$$|\tilde{c}(4K_0) - 4\tilde{c}(K_0) - 12k^{-1}K_0^2| < \epsilon^2.$$

Furthermore, by the definition of K_0 and (8.5), it follows that

$$\max(12k^{-1}K_0^2, 12|A_- - A_+| \frac{K_0 \log K_0}{\ell \log \ell}) \geq \frac{1}{1000\bar{A}^2}.$$

¹⁰In what follows we will often use the following property of $\tilde{b}(\cdot)$: for an interval $[R, S] \subset [0, \tilde{T}]$ such that $S \leq R^{1+\epsilon^9}$ and every $t \in [R, S]$, we have

$$|\tilde{b}(t) - (A_- - A_+) \frac{t \log R}{\ell \log \ell}| \leq \epsilon^5 |A_- - A_+| \frac{t \log R}{\ell \log \ell} \leq \epsilon^3,$$

since $t, R \leq \tilde{T}$. This means that for such $[R, S]$, the function $\tilde{b}(\cdot)$ is almost linear with coefficient $(A_- - A_+) \frac{\log R}{\ell \log \ell}$.

The three above equations imply that $\max(|\tilde{c}(4K_0)|, |\tilde{c}(K_0)|) > \frac{1}{32 \cdot 2018 A^2}$. This finishes the proof of the claim (8.13).

Let $u := \frac{1}{32 \cdot 2018 A^2}$. Notice that by (8.5) and the definition of u , we have

$$\tilde{c}(u^2 \tilde{T}) \leq k^{-1} \left(u^2 \tilde{T} \right)^2 + |A_- - A_+| \frac{u^2 \tilde{T} \log u^2 \tilde{T}}{\ell \log \ell} \leq u.$$

Therefore and since $\tilde{c}(\cdot)$ is continuous, by (8.13), we have $|\tilde{c}(T_1)| = u$ for some $T_1 \in [u^2 \tilde{T}, \frac{\tilde{T}}{2A}]$. Moreover, for every $t \in [T_1, (1 + \kappa)T_1]$, we have (since $T_1 \leq \tilde{T}$ and by (8.5))

$$(8.14) \quad |\tilde{c}(t) - \tilde{c}(T_1)| \leq k^{-1}(t^2 - T_1^2) + \frac{|A_- - A_+|}{\ell \log \ell} (t \log t - T_1 \log T_1) \leq k^{-1}(2\kappa \tilde{T} + \kappa^2 \tilde{T}^2) + \frac{2\kappa |A_- - A_+| \tilde{T} \log \tilde{T}}{\ell \log \ell} \leq \epsilon^3,$$

where the last inequality follows by the definition of κ .

We set $M := T_1$ and $L := \kappa T_1$. Let

$$(8.15) \quad U := \left\{ t \in [M, M + L] : (R_\alpha)_t^f(y, s) \in \{(y, s) \in \mathbb{T}^f : \epsilon^2 < s < f(y) - \epsilon^2\} \right\}.$$

By (8.1), it follows that $|U| \geq (1 - \epsilon^{3/2})|L|$. Moreover, by (8.15), we have $U = \bigcup_{i=1}^w (c_i, d_i)$, where $|d_i - c_i| \geq \inf_{\mathbb{T}} f - 2\epsilon^2$ for every $i = 2, \dots, w-1$ and hence $w \leq \frac{|L|}{\bar{d}}$, where $\bar{d} = \frac{\inf_{\mathbb{T}} f}{2}$. Let $a(t) := t - \tilde{b}(c_i)$ on (c_i, d_i) , $i = 1, \dots, w$ and we set $a(t) = 0$ for $t \in [M, M + L] \setminus U$. Notice that by (8.5) (since $c_1 \leq \tilde{T}$), we have

$$(8.16) \quad |\tilde{b}(c_1)| = |A_- - A_+| \frac{c_1 \log c_1}{\ell \log \ell} \leq |A_- - A_+| \leq \epsilon^2 L,$$

since $\epsilon^2 L = \epsilon^2 \kappa T_1 \geq \kappa^3 \tilde{T} \geq \kappa^4 \delta^{-1/2} > \kappa^{-1} > |A_- - A_+|$. Moreover, by (7.10), for $i = 1, \dots, w$, $|c_{i+1} - c_i| \leq |f^{(n(y, s, c_i))}(y)| + 2\epsilon^2 = O(\log q_v) < \epsilon^3 q_v < \epsilon^3 \ell$ and hence, we have (since $c_{i+1} \leq \tilde{T} \leq \ell$)

$$(8.17) \quad |\tilde{b}(c_{i+1}) - \tilde{b}(c_i)| \leq 2|A_- - A_+| \frac{(c_{i+1} - c_i) \log c_{i+1}}{\ell \log \ell} \leq \epsilon^3$$

since $|c_{i+1} - c_i| \leq \epsilon^3 \ell$ and $c_{i+1} \leq \tilde{T} \leq \ell$ (see (8.5)) for $i = 1, \dots, w$. Therefore, by (8.16), (8.17) and the definition of U , (a, U, \bar{d}) is $\epsilon^{3/2}$ -good (see (PAL) in Definition 3.1). Moreover, for $t \in (c_i, d_i)$, $i = 1, \dots, w$, by (8.17), we have

$$(8.18) \quad |a(t) - (t - \tilde{b}(t))| = |-\tilde{b}(c_i) + \tilde{b}(t)| \leq |\tilde{b}(c_{i+1}) - \tilde{b}(c_i)| \leq \epsilon^3.$$

By (8.12), (8.15) and (8.18), since $[M, M + L] \subset [0, \tilde{T}]$, we have

$$(8.19) \quad d^f((R_\alpha)_t^f((y, s)), (R_\alpha)_{a(t)}^f((y', s'))) < \epsilon \text{ for every } t \in U \cap [M, M + L].$$

Furthermore, by (8.18), we have $t - k^{-1}t^2 := (t - \tilde{b}(t)) + \tilde{c}(t) = a(t) + \tilde{c}(t) + O(\epsilon^3)$. Since $\tilde{c}(M) = \pm u = \pm \frac{1}{32 \cdot 2018 A^2} \in P$, by (8.6) and (8.14), we have

$$(8.20) \quad d_{G/\Gamma}(\tilde{h}_t x, \tilde{h}_{a(t)+p} x') < \epsilon \text{ for every } t \in [M, M + L].$$

By (8.19) and (8.20), it follows that (3.2) holds and this, by Theorem 3.1, finishes the proof of Theorem 1.3.

9. TIME CHANGES OF HOROCYCLE FLOWS AND SARNAK'S CONJECTURE
(ANSWER TO M. RATNER'S QUESTION)

We now turn our attention to the problem of Möbius disjointness of time automorphisms of flows considered in the present paper. Recall that the Möbius function $\mu : \mathbb{N} \rightarrow \{-1, 0, 1\}$ is defined as $\mu(1) = 1$, $\mu(p_1 \dots p_k) = (-1)^k$ for pairwise different prime numbers p_1, \dots, p_k and $\mu(n) = 0$ for the remaining $n \in \mathbb{N}$. It is not hard to see that $\mu(mn) = \mu(m)\mu(n)$ whenever m and n are coprime, i.e. μ is an example of arithmetic *multiplicative* function. Hence, μ is a member of

$$\mathcal{M}_1 := \{\mathbf{u} : \mathbb{N} \rightarrow \mathbb{C} : \mathbf{u} \text{ is multiplicative and } |\mathbf{u}| \leq 1\}.$$

A basic method to prove disjointness of a bounded numeric sequence $(a_n) \subset \mathbb{C}$ with all members of \mathcal{M}_1 (in fact, with all bounded multiplicative functions) is the following criterion:

Proposition 9.1 ([22],[5]). *If for all sufficiently large prime numbers $p \neq q$ we have*

$$(9.1) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} a_{pn} \bar{a}_{qn} = 0,$$

then

$$(9.2) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} a_n \mathbf{u}(n) = 0$$

for each $\mathbf{u} \in \mathcal{M}_1$.

Given a topological dynamical system (X, T) , where T is a homeomorphism of a compact metric space X , we want to study *disjointness* of (X, T) with $\mathbf{u} \in \mathcal{M}_1$, meaning that $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} f(T^n x) \mathbf{u}(n) = 0$ for each $f \in C(X)$ and $x \in X$. As constant functions are continuous, to have such a disjointness, we need to assume that, additionally to $\mathbf{u} \in \mathcal{M}_1$, the mean of \mathbf{u} exists and equals zero:

$$(9.3) \quad M(\mathbf{u}) := \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} \mathbf{u}(n) = 0$$

(μ has mean zero, by the Prime Number Theorem).

In what follows, we consider the problem of disjointness of (X, T) only with the members in \mathcal{M}_1 satisfying (9.3).

Now, assume additionally that (X, T) is a uniquely ergodic topological dynamical system, with the unique T -invariant measure μ which makes (X, μ, T) totally ergodic. If we fix $x \in X$ then each accumulation point ρ of the empiric measures $\frac{1}{N} \sum_{n \leq N} \delta_{(T^p \times T^q)^n(x, x)}$, $N \geq 1$, yields a member of $J(T^p, T^q)$ (indeed, T^p and T^q are also uniquely ergodic, and μ is their only invariant measure). If we take a zero mean $f \in C(X)$ and $T^p \perp T^q$ (for $p \neq q$ sufficiently large), then immediately, $T^p \times T^q$ is still uniquely ergodic (with $\mu \otimes \mu$ being the unique invariant measure) and we find that the sequence (a_n) , $a_n = f(T^n x)$, satisfies (9.1) (indeed, the limit equals $\int_{X \times X} f \otimes \bar{f} d(\mu \otimes \mu) = 0$). Hence, cf. [5], if the different, sufficiently large, prime powers of T are disjoint then we find that the system (X, T) is orthogonal to any $\mathbf{u} \in \mathcal{M}_1$ satisfying (9.3).

Notice that the flows (T_t) we are considering are mixing. Hence, all non-zero time automorphisms T_{t_0} are totally ergodic. When considering any uniquely ergodic model of T_{t_0} (in fact, those considered in Theorem 1.1 are themselves uniquely ergodic), they satisfy the disjointness assumption: $T_{pt_0} \perp T_{qt_0}$ for $p \neq q$ prime numbers sufficiently large, whence the assumption (9.1) is satisfied for each x and zero mean f (in the model) and we obtain

disjointness with \mathbf{u} . As a matter of fact, the result holds also for $t_0 = 0$ as the mean of \mathbf{u} equals 0.

On the other hand, to distinguish between Möbius disjointness of zero entropy systems and positive entropy systems¹¹, it is proved in [2] that the following conditions are equivalent (for the definition of strong \mathbf{u} -MOMO, see Appendix C):

Proposition 9.2. *The following conditions are equivalent:*

- (i) *Sarnak's conjecture holds.*
- (ii) *Sarnak's conjecture holds uniformly (in $x \in X$).*
- (iii) *Strong μ -MOMO property holds for each zero (topological) entropy system.*

Motivated by some recent break-through results in multiplicative number theory [26], [27] and the role of convergence on (typical) short intervals (see [11] for more details), we will now consider a fourth natural condition of orthogonality which is a uniform short interval convergence (USIC for short) in (X, T) :

Definition 9.1. (X, T) satisfies the *strong \mathbf{u} -USIC¹² property* if

$$\frac{1}{M} \sum_{M \leq m < 2M} \left| \frac{1}{H} \sum_{m \leq h < m+H} f(T^h x) \mathbf{u}(h) \right| \rightarrow 0$$

when $H, M \rightarrow \infty$, $H = o(M)$, uniformly in $x \in X$.

The meaning of convergence in Definition 9.1 is the following: For all sequences $(M_\ell), (H_\ell)$ tending to infinity with $H_\ell/M_\ell \rightarrow 0$ when $\ell \rightarrow \infty$, we have

$$(9.4) \quad \lim_{\ell \rightarrow \infty} \frac{1}{M_\ell} \sum_{M_\ell \leq m < 2M_\ell} \left| \frac{1}{H_\ell} \sum_{m \leq h < m+H_\ell} f(T^h x) \mathbf{u}(h) \right| = 0$$

uniformly in $x \in X$.

Note also that by considering $M, 2M, \dots, 2^k M, \dots$, both in Definition 9.1 and in (9.4) we can consider (and we will) the sum from 1 to M instead of the sum from M to $2M$.

In Appendix C we show that a system (X, T) satisfies the strong \mathbf{u} -MOMO property if and only if it satisfies the strong \mathbf{u} -USIC property.

Corollary 9.1. *The following condition*

- (iv) *μ -USIC property holds for each zero (topological) entropy system is equivalent to (i) in Proposition 9.2 (and hence to (ii) and (iii)).*

Question It is proved in [2] that the systems satisfying so called AOP property, in particular systems whose prime powers are disjoint, satisfy the strong \mathbf{u} -MOMO property. Collecting all the results, we obtain the following.

Corollary 9.2. *For each flow (T_t) being either a uniquely ergodic model of an Arnol'd flow satisfying the assumptions of Theorem 1.2 or a non-trivial smooth time-change in $B^+(M)$*

¹¹As proved by Downarowicz and Serafin [9], there are positive entropy systems which are Möbius disjoint. On the other hand no positive entropy system satisfies the strong μ -MOMO property [2].

¹²Acronym for Uniform Short Intervals Convergence. Note that the strong \mathbf{u} -USIC property requires for \mathbf{u} to satisfy convergence on (typical) short intervals which we obtain by taking $f = 1$ in Definition 9.1. In view of [26], μ fulfills this requirement. In what follows, we consider only \mathbf{u} satisfying the relevant short interval convergence.

of a horocycle flow in the cocompact case, we obtain the following: For each $t_0 \in \mathbb{R}$ and each $f \in C(X)$, we have

$$\frac{1}{M} \sum_{1 \leq m < M} \left| \frac{1}{H} \sum_{m \leq h < m+H} f(T_{ht_0}x) \mathbf{u}(h) \right| \rightarrow 0$$

when $H, M \rightarrow \infty$, $H = o(M)$, uniformly in $x \in X$.

In particular, for such flows, Sarnak's conjecture holds uniformly.

Corollary 9.2 brings the positive answer to M. Ratner's question (see 7 in [11]). Although, Möbius disjointness itself is known for horocycle flows [5], it remains however an open question whether the assertions of Corollary 9.2 hold for horocycle flows themselves even when $\mathbf{u} = \boldsymbol{\mu}$.

Remark 9.1. If we have a dynamical system (X, T) for which the sums

$$(*) \quad \frac{1}{N} \sum_{n \leq N} h(T^n x) \boldsymbol{\mu}(n)$$

converge to zero uniformly in $x \in X$ (for each $h \in C(X)$) then one can show (considering respectively the sums $(*)$ with $T^{-N}x$ and T^{-1} or with $T^N x$ and T) that we have also convergence for sums of the following type:

$$(9.5) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} h(T^n x) \boldsymbol{\mu}(N - n) = 0,$$

$$(9.6) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} h(T^n x) \boldsymbol{\mu}(N + n) = 0.$$

We do not know if (9.5), (9.6) hold in case of horocycle flows or locally Hamiltonian flows on \mathbb{T}^2 .

APPENDIX A. CONSEQUENCES OF SHEARING FOR TIME CHANGES OF HOROCYCLE FLOWS

This appendix is devoted to the proof of Proposition 4.1. We use the notation introduced in Section 4.1

Proof of Proposition 4.1. We will for simplicity assume that $\int_M \tau d\mu = 1$. Fix $0 < \epsilon < K^{-3}$. Using the unique ergodicity of (h_t) , let $N_\epsilon > 0$ be such that for every $T \in \mathbb{R}$ satisfying $|T| > N_\epsilon$ and every $y \in M$, remembering that $\int_M X\tau d\mu = 0$, we have

$$(A.1) \quad \max \left(\left| \int_0^T \tau(h_t y) dt - T \right|, \left| \int_0^T (X\tau)(h_t y) dt \right| \right) \leq \epsilon^3 |T|$$

and (cf. Remark 2.1)

$$(A.2) \quad |u(T, y) - T| \leq \epsilon^4 |T|.$$

Moreover, let $\delta := \min(\epsilon^{20}, N_\epsilon^{-20})$. Take $x, y \in M$ with $\max(|r|, |s|, |\bar{v}|) < \delta$ and let $T \in \mathbb{R}$ be such that $|T| \in [N_\epsilon, K|r|^{-1/2}]$. We will additionally assume that $T > 0$, the case $T < 0$ is analogous.

We have

$$(A.3) \quad h_T x y^{-1} h_{-\chi_{x,y}(T)} = h_{\bar{v}} \begin{pmatrix} e^s + rT & v(x, y, T) \\ r & e^{-s} - r\chi_{x,y}(T) \end{pmatrix},$$

where (using also (4.2))

$$v(x, y, T) = Te^{-s} - \chi_{x,y}(T)e^s - T\chi_{x,y}(T)r = e^{-3s}r^2T^3.$$

By the definition of δ , we hence obtain (remembering that $|r| < \delta$)

$$(A.4) \quad |v(x, y, T)| = e^{-3s}|r|^2T^3 \leq 2|r|^2K|r|^{-3/2} \leq 2K\delta^{1/2} < \epsilon^2.$$

In what follows (since x, y are now fixed), we will write χ instead of $\chi_{x,y}$.

By (2.4) and (4.3), we obtain

$$A_x(T) = \int_{u(\chi(T), y)}^{u(\chi(T) + A_x(T), y)} \tau(h_\theta y) d\theta = \int_{u(\chi(T), y)}^{\chi(u(T, x))} \tau(h_\theta y) d\theta.$$

Representing the last term as the difference of two integrals, where $\int_0^{u(\chi(T), y)} \tau(h_\theta y) d\theta = \chi(T)$, the assertion (4.4) can be rewritten as

$$(A.5) \quad \left| \int_0^{\chi(u(T, x))} \tau(h_t y) dt - \chi(T) + e^{-2s} \int_0^{u(T, x)} (\tau - \tau \circ g_{-s})(h_t x) dt \right| \leq \epsilon.$$

Changing variables (t replaced by $\chi(\theta)$), we have

$$(A.6) \quad \begin{aligned} \int_0^{\chi(u(T, x))} \tau(h_t y) dt - \chi(T) &= \int_0^{u(T, x)} \tau(h_{\chi(t)} y) \chi'(t) dt - \chi(T) = \\ &= e^{-2s} \int_0^{u(T, x)} \tau(h_t x) dt + e^{-2s} \left(\int_0^{u(T, x)} (-\tau(h_t x) + \tau(h_{\chi(t)} y) dt) \right) + \\ &\quad \int_0^{u(T, x)} \tau(h_{\chi(t)} y) (\chi'(t) - e^{-2s}) dt - \chi(T). \end{aligned}$$

We now claim that

$$(A.7) \quad \left| \int_0^{u(T, x)} (-\tau(h_t x) + \tau(h_{\chi(t)} y)) dt + \int_0^{u(T, x)} (\tau - \tau \circ g_{-s})(h_t x) dt \right| \leq \epsilon^{3/2}.$$

Indeed, this claim is equivalent to showing

$$(A.8) \quad \left| \int_0^{u(T, x)} (\tau(h_{\chi(t)} y) - \tau(g_{-s} h_t x)) dt \right| \leq \epsilon^{3/2}.$$

Notice that by (A.3), we have

$$(A.9) \quad \begin{aligned} g_{-s} h_t x y^{-1} h_{-\chi(t)} &= h_{\bar{v}e^{-2s}} \begin{pmatrix} 1 + e^{-s} r t & e^{-s} v(x, y, t) \\ e^s r & 1 - e^s r \chi(t) \end{pmatrix} = \\ &= h_{w(t) + \bar{v}e^{-2s}} \begin{pmatrix} (1 - e^s r \chi(t))^{-1} & 0 \\ e^s r & 1 - e^s r \chi(t) \end{pmatrix}, \end{aligned}$$

where $w(t) = \frac{e^{-s} v(x, y, t)}{1 - e^s r \chi(t)}$ (recall that $v(x, y, t) = e^{-3s} r^2 t^3$). We also have $v'(x, y, t) = 3e^{-3s} r^2 t^2 \leq C'|r|$ for some constant $C' > 0$ (since $t \leq K|r|^{-1/2}$). Moreover, $|r\chi(t)| = O(|r|t + r^2 t^2) = O(r^{1/2})$ and

$$|v(x, y, t) r \chi'(t)| = O(r^2 t^3 |r| (1 + |r|t)) = O(|r|^{3/2}).$$

Therefore (enlarging C' if necessary), we have

$$(A.10) \quad |w'(t)| \leq 2C'|r|.$$

Now, (A.8) (in particular) follows by showing that for every $S \in [0, T]$ and $\Psi \in \{\tau, X\tau\}$, we have

$$(A.11) \quad \left| \int_0^{u(S,x)} (\Psi(h_{-w(t)-\bar{v}e^{-2s}}g_{-s}h_tx) - \Psi(g_{-s}h_tx)) dt \right| < \frac{\epsilon^3}{2}$$

and

$$(A.12) \quad \left| \int_0^{u(T,x)} (\tau(h_{\chi(t)}y) - \tau(h_{-w(t)-\bar{v}e^{-2s}}g_{-s}h_tx)) dt \right| < \frac{\epsilon^{3/2}}{2}.$$

We will first show (A.11). Using $g_{-s}h_t = h_{e^{-2s}t}g_{-s}$, substituting $t' = e^{-2s}t$ and making use of (A.10), we get

$$\begin{aligned} \int_0^{u(S,x)} \Psi(h_{-w(t)-\bar{v}e^{-2s}}g_{-s}h_t) dt &= e^{2s} \int_0^{u(e^{-2s}S,x)} \Psi(h_{t'-w(e^{2s}t')-\bar{v}e^{-2s}}g_{-s}x) dt' = \\ &= e^{2s} \int_0^{u(e^{-2s}S,x)} \Psi(h_{t-w(e^{2s}t)-\bar{v}e^{-2s}}g_{-s}x)(1 - e^{2s}w'(e^{2s}t)) dt + O(|r|S). \end{aligned}$$

Substituting $t' = t - w(e^{2s}t) - \bar{v}e^{-2s}$ and using $\max(w(0) + \bar{v}e^{-2s}, w(e^{2s}u(e^{-2s}S, x)) - \bar{v}e^{-2s}) < \delta^{1/4}$, we get

$$\begin{aligned} e^{2s} \int_0^{u(e^{-2s}S,x)} \Psi(h_{t-w(e^{2s}t)-\bar{v}e^{-2s}}g_{-s}x)(1 - e^{2s}w'(e^{2s}t)) dt &= \\ e^{2s} \int_0^{u(e^{-2s}S,x)} \Psi(h_{t'}g_{-s}x) dt' + O(\delta^{1/4}) &= \int_0^{u(S,x)} \Psi(g_{-s}h_tx) dt + O(\delta^{1/4}), \end{aligned}$$

where the latter equality follows by substituting $t = e^{2s}t'$ and using the renormalization identity. This finishes the proof of (A.11).

We will now show (A.12). To begin the proof of that claim, note that by (A.9), we have

$$h_{-w(t)-\bar{v}e^{-2s}}g_{-s}h_tx = \begin{pmatrix} (1 - e^s r \chi(t))^{-1} & 0 \\ e^s r & 1 - e^s r \chi(t) \end{pmatrix} h_{\chi(t)}y,$$

whence $d_U(h_{-w(t)-\bar{v}e^{-2s}}g_{-s}h_tx, h_{\chi(t)}y) = 0$. Now, by Lemma 4.1, we get that the LHS of (A.12) is equal to

$$\begin{aligned} &\int_0^{u(T,x)} \sum_{W \in \{X, V\}} (W\tau)(h_{-w(t)-\bar{v}e^{-2s}}g_{-s}h_tx) d_W(h_{-w(t)-\bar{v}e^{-2s}}g_{-s}h_tx, h_{\chi(t)}y) dt + \\ &O \left(\epsilon^3 u(T, x) \sup_{t \in [0, u(T, x)]} \left| \sum_{W \in \{X, V\}} d_W(h_{-w(t)-\bar{v}e^{-2s}}g_{-s}h_tx, h_{\chi(t)}y) \right| \right). \end{aligned}$$

Moreover, by (A.9), Remark 4.1 and (A.2), the above expression is equal to

$$\int_0^{u(T,x)} (X\tau)(h_{-w(t)-\bar{v}e^{-2s}}g_{-s}h_tx) e^s r \chi(t) dt + \int_0^{u(T,x)} (V\tau)(h_{-w(t)-\bar{v}e^{-2s}}g_{-s}h_tx) e^s r dt + O(\epsilon^3 T^2 |r|).$$

The middle summand we estimate by $O(e^s |r| u(x, T))$ which by (A.2) and the fact that $|r| < \delta^{1/2} < \epsilon^3$ is $O(\epsilon^3)$ and since $T \leq K|r|^{-1/2}$, the last expression is equal to

$$r e^s \int_0^{u(T,x)} \chi(t) (X\tau)(h_{-w(t)-\bar{v}e^{-2s}}g_{-s}h_tx) dt + O(\epsilon^3).$$

Since $re^s\chi(t) = e^{-s}rt + e^{-2s}r^2t^2 = e^{-s}rt + O(r)$, and $u(T, x) \leq 2T \leq 2Kr^{-1/2}$ (see (A.2)), in order to show (A.12), it is enough to establish

$$(A.13) \quad \left| \int_0^{u(T,x)} t(X\tau)(h_{-w(t)-\bar{v}e^{-2s}}g_{-s}h_tx) dt \right| = O(\epsilon^2|r|^{-1}).$$

Integration by parts yields

$$\begin{aligned} & \int_0^{u(T,x)} t(X\tau)(h_{-w(t)-\bar{v}e^{-2s}}g_{-s}h_tx) dt = \\ & T \int_0^{u(T,x)} (X\tau)(h_{-w(t)-\bar{v}e^{-2s}}g_{-s}h_tx) dt - \int_0^{u(T,x)} \left(\int_0^t (X\tau)(h_{-w(\theta)-\bar{v}e^{-2s}}g_{-s}h_\theta x) d\theta \right) dt. \end{aligned}$$

Therefore and by (A.11) (for $\Psi = X\tau$ and $S = u(T, x)$ and then $S = t$), (A.13) follows by showing that

$$\left| \int_0^t (X\tau)(g_{-s}h_\theta x) d\theta \right| = O(\epsilon^3|r|^{-1/2}),^{13}$$

this however, by the renormalization equation and substituting $t' = e^{-2s}\theta$, follows by showing that

$$\left| \int_0^{e^{-2s}t} (X\tau)(h_{t'}g_{-s}x) dt' \right| = O(\epsilon^3|r|^{-1/2}),$$

which is true by (A.1) if $e^{-2s}t > N_\epsilon$ (if not, we estimate $O(N_\epsilon) = O(\epsilon^3|r|^{-1/2})$). This finishes the proof of (A.13) and hence also the proof of (A.12). Now, the proof of (A.8) is complete and the claim (A.7) follows.

Next, by definition, $e^{-2s} \int_0^{u(T,x)} \tau(h_tx)dt - \chi(T) = e^{-3s}rT^2$, so to finish the proof of (A.5), by (A.6) and (A.7), it is enough to show that

$$(A.14) \quad \left| \int_0^{u(T,x)} \tau(h_{\chi(t)}y)(\chi'(t) - e^{-2s}) dt + e^{-3s}rT^2 \right| < \epsilon^{5/4}.$$

Let us notice that we consider the function $\chi(t) = e^{-2s}t - e^{-3s}rt^2$ on the interval $[0, O(r^{1/2})]$ and that its maximal value is obtained at $e^s/2r$. It follows that χ on the interval under consideration is invertible and we denote the inverse function by $\chi^{-1}(w) = t$. Set $m(w) := 1 - \frac{e^{-2s}}{\chi'(\chi^{-1}(w))}$. Moreover, notice that $m(w) = m(\chi(t)) = \frac{-2e^{-3s}rt}{e^{-2s}-2e^{-3s}rt}$. By a direct computation, we verify the following properties of m :

- (z1) $|wm(w)| = O(1)$, for every $\chi^{-1}(w) \in [0, O(|r|^{-1/2})]$; indeed, $wm(w) = \chi(t) \frac{-2e^{-3s}rt}{e^{-2s}-2e^{-3s}rt} = O(rt^2) = O(1)$.
- (z2) $|m(N_\epsilon)| < \delta^{1/2}$; indeed, $N_\epsilon \in [\chi(e^{2s}N_\epsilon - 1), \chi(e^{2s}N_\epsilon + 1)]$
- (z3) $|m'(w)| > 0$ for every $\chi^{-1}(w) \in [0, O(|r|^{-1/2})]$; indeed, $m'(w) = m'(\chi(t))\chi'(t)$ and both terms in the product are of constant sign.

Changing variable ($w = \chi(t)$) and then integrating by parts, we have

$$(A.15) \quad \begin{aligned} & \int_0^{u(T,x)} \tau(h_{\chi(t)}y)(\chi'(t) - e^{-2s}) dt = \int_0^{\chi(u(T,x))} \tau(h_wy)m(w) dw = \\ & m(\chi(u(T,x))) \int_0^{\chi(u(T,x))} \tau(h_wy) dw - \int_0^{\chi(u(T,x))} \left(\int_0^w \tau(h_\theta y) d\theta \right) m'(w) dw. \end{aligned}$$

¹³Indeed, the first expression is then multiplied by T and the second by evaluation of the integrand yields multiplication by $u(T, x)$, in both cases, we obtain $O(\epsilon^3|r|^{-1/2}T) = O(\epsilon^3|r|^{-1})$.

Therefore, by (A.1) and (z1), we have

$$(A.16) \quad \left| m(\chi(u(T, x))) \int_0^{\chi(u(T, x))} \tau(h_w y) dw - m(\chi(u(T, x))) \chi(u(T, x)) \right| = O(\epsilon^3).$$

Moreover,

$$(A.17) \quad \int_0^{\chi(u(T, x))} \left(\int_0^w \tau(h_\theta y) d\theta \right) m'(w) dw = \int_0^{\chi(u(T, x))} (w + p(w)) m'(w) dw,$$

where, by (A.1), $|p(w)| < \epsilon^2 w$ for $w \geq N_\epsilon$, and $|p(w)| = O(N_\epsilon)$ for $w \in [0, N_\epsilon]$. Notice that by integrating by parts, we have

$$(A.18) \quad \int_0^{\chi(u(T, x))} w m'(w) dw = m(\chi(u(T, x))) \chi(u(T, x)) - \int_0^{\chi(u(T, x))} m(w) dw.$$

But changing variables, $w = \chi(t)$, we obtain

$$(A.19) \quad \int_0^{\chi(u(T, x))} m(w) dw = \int_0^{u(T, x)} (\chi'(t) - e^{-2s}) dt = \chi(u(T, x)) - e^{-2s} u(T, x) = -e^{-3s} r(u(T, x))^2.$$

In particular, from (A.18), (A.19) and (z1), we obtain

$$(A.20) \quad \int_0^{\chi(u(T, x))} w m'(w) dw = O(1).$$

By (z3), (z2), (A.2) and (A.20), we get

$$(A.21) \quad \left| \int_0^{\chi(u(T, x))} p(w) m'(w) dw \right| \leq \left| \int_0^{N_\epsilon} p(w) m'(w) dw \right| + \left| \int_{N_\epsilon}^{\chi(u(T, x))} p(w) m'(w) dw \right| \leq O(N_\epsilon) m(N_\epsilon) + \epsilon^2 \left| \int_0^{\chi(u(T, x))} w m'(w) dw \right| = O(\epsilon^2).$$

Finally, in view of (A.15), (A.16), (A.17), (A.18), (A.19) and (A.21), we get

$$\left| \int_0^{u(T, x)} \tau(h_{\chi(t)} y) (\chi'(t) - e^{-2s}) dt + e^{-3s} r(u(T, x))^2 \right| = O(\epsilon^2).$$

Hence (A.14) follows by the triangle inequality since by (A.2) and $T \leq K r^{-1/2}$, we have

$$|e^{-3s} r(u(T, x))^2 - e^{-3s} r T^2| \leq e^{-3s} |r| (u(T, x)^2 - T^2) \leq \epsilon^2 e^{-3s} r T^2 \leq \epsilon^{3/2}.$$

This finishes the proof of (4.4).

For the proof of (4.5), notice that $\tilde{h}_T(x) = h_{u(T, x)}(x)$ and, by (4.3), $\tilde{h}_{\chi(T) + A_x(T)} y = h_{u(\chi(T) + A_x(T), y)} y = h_{\chi(u(T, x))} y$. The statement now follows by (A.3) and (A.4), where T is replaced by $u(T, x)$. \square

APPENDIX B. ERGODIC AVERAGES FOR HOROCYCLE FLOWS

The following result (Lemma B.1 below) on ergodic averages for horocycle flows follows from the work of Flaminio-Forni [12] and Bufetov-Forni [6]. We are indebted to Giovanni Forni for explaining it to us.

For any smooth real-valued function $f \in W^r(M)$, with $r > 11/2$, let

$$f = \sum_{\mu \in \text{spec}(\square)} f_\mu$$

denote the decomposition of f with respect to a splitting of the space $L^2(M)$ into irreducible components, parametrized by the eigenvalues μ of the Casimir operator \square (listed with multiplicities), cf. Section 4.2. That is, f_μ is the projection of f on the irreducible component of Casimir parameter μ . Set

$$\mu_f := \min\{\mu \in \text{spec}(\square) \setminus \{0\} : f_\mu \neq 0\}.$$

Lemma B.1 (ergodic averages estimates). *For any real-valued function $\tau \in W^r(M)$ (for any $r > 11/2$) with non-trivial support on the irreducible components of the complementary series there exist a function $\phi_\tau \in C^\infty(\mathbb{R})$, with $\phi_\tau(0) = 0$ and $\phi'_\tau(0) \neq 0$, and a function $\beta_\tau \in C(M)$ such that the following holds. There exist constants $C_r > 0$, $\alpha_\tau \in (0, 1)$ and $\gamma_\tau \in (0, \alpha_\tau)$ such that, for every $(x, T) \in M \times \mathbb{R}$, we have*

$$|T^{-\alpha_\tau} \int_0^T (\tau - \tau \circ g_s) \circ h_t(x) dt - \phi_\tau(s) \beta_\tau(g_{\log T} x)| \leq C_r s \|\tau\|_r T^{-\gamma_\tau}.$$

For any function $\tau \in W^r(M)$ (for any $r > 11/2$) with trivial support on the irreducible components of the complementary series, and not fully supported on the discrete series, there exist functions $\beta_\tau \in C^\infty(\mathbb{R} \times \mathbb{T}^\infty, C(M))$ and $\beta_\tau^{(1/4)} \in C(M)$ with $\beta_\tau(0, \theta, x) \equiv 0$ and $\frac{d}{ds} \beta_\tau(0, \theta, x) \not\equiv 0$, and a vector $\mathbf{v} \in \mathbb{T}^\infty$, such that the following holds. There exist constants $C_r > 0$ and $\gamma \in (0, 1/2)$ such that, for every $(x, T) \in M \times \mathbb{R}$, we have

$$\begin{aligned} |T^{-\frac{1}{2}} \int_0^T (\tau - \tau \circ g_s) \circ h_t(x) dt - \beta_\tau(s, \mathbf{v} \log T, g_{\log T} x) \\ - (e^{-s/2} - 1) \beta_\tau^{(1/4)}(g_{\log T} x) \log T| \leq C_r s \|\tau\|_r T^{-\gamma}. \end{aligned}$$

The function $\beta_\tau^{(1/4)}$ vanishes identically if and only if the projection $\tau_{1/4}$ of the function $\tau \in W^r(M)$ on the irreducible component of Casimir parameter $1/4$ is a coboundary of the horocycle flow.

Proof. We distinguish three cases: $\mu_f < 1/4$, $\mu_f > 1/4$ and $\mu_f = 1/4$. In the first case, $\mu_f < 1/4$, let H_1, \dots, H_k denote all the irreducible components of Casimir parameters $\mu_1 = \dots = \mu_k = \mu_f$. Let D_1^\pm, \dots, D_k^\pm denote the basis of distributional eigenvectors of the geodesic flow of the space of invariant distributions for the horocycle flow supported on $\mathcal{D}'(H_1) \oplus \dots \oplus \mathcal{D}'(H_k)$. Let $\beta_1^\pm, \dots, \beta_k^\pm$ be the corresponding cocycles for the horocycle flow. In this case, the components H_1, \dots, H_k belong to the complementary series and by [6], Lemma 3.1, the following holds. Let $\nu_f := (1 - 4\mu_f)^{1/2} \in \mathbb{R}^+$. There exist constants $C_r > 0$ and $\gamma_f \in (0, \frac{1+\nu_f}{2})$ such that, for all $(x, T) \in M \times \mathbb{R}$ we have

$$(B.1) \quad |T^{-\frac{1+\nu_f}{2}} \int_0^T f \circ h_t(x) dt - \text{Re} \sum_{i=1}^k D_i^-(f) \beta_i^-(g_{\log T} x, 1)| \leq C_r \|f\|_r T^{-\gamma_f};$$

in the second case, $\mu_f > 1/4$, the function f has no components on the complementary series and no component corresponding to the Casimir eigenvalue $\mu = 1/4$. Let then $(\mu_n)_{n \in \mathbb{N}}$ denote the sequence of Casimir parameters in the interval $(1/4, +\infty)$ (listed with multiplicities), let $(D_{\mu_n}^\pm)$ denote the sequence of normalized horocycle invariant distributions and let $(\beta_{\mu_n}^\pm)$ denote the corresponding sequence of additive Hölder cocycles. By [6], Lemma 3.2, the following holds:

For every $n \in \mathbb{N}$, let $v_n = (4\mu_n - 1)^{1/2} \in \mathbb{R}^+$. There exist constants $C_r > 0$ and $\gamma \in (0, 1/2)$ such that, for all $(x, T) \in M \times \mathbb{R}$,

$$(B.2) \quad \left| T^{-1/2} \int_0^T f \circ h_t(x) dt - \operatorname{Re} \sum_{n \in \mathbb{N}} D_{\mu_n}^+(f) \exp(i \frac{v_n}{2} \log T) \beta_{\mu_n}^+(g_{\log T} x, 1) \right| \leq C_r \|f\|_r T^{-\gamma}.$$

Finally, in the case $\mu_f = 1/4$, we have to add to the above expansion for the case $\mu > 1/4$ the following contribution of the irreducible components with Casimir parameter $\mu = 1/4$. Let $H_{1/4,1}, \dots, H_{1/4,l}$ denote all the irreducible components of Casimir parameters $\mu_1 = \dots = \mu_l = 1/4$. Let $D_{1/4,1}^\pm, \dots, D_{1/4,l}^\pm$ denote the basis of distributional eigenvectors of the geodesic flow of the space of invariant distributions for the horocycle flow supported on $\mathcal{D}'(H_{1/4,1}) \oplus \dots \oplus \mathcal{D}'(H_{1/4,l})$. Let $\beta_{1/4,1}^\pm, \dots, \beta_{1/4,l}^\pm$ be the corresponding cocycles for the horocycle flow. By the exact scaling and asymptotic results of [6], Theorem 1.2 and Corollary 1.3, or Corollary 3.2, it follows that for the projection $f_{1/4}$ of $f \in W^r(M)$ onto the component $H_{1/4,1} \oplus \dots \oplus H_{1/4,l}$ of the Hilbert space $W^r(M)$, the following holds. There exist constants $C_r > 0$ and $\gamma \in (0, 1/2)$ such that, for all $(x, T) \in M \times \mathbb{R}$,

$$(B.3) \quad \left| T^{-1/2} \int_0^T f_{1/4} \circ h_t(x) dt - \sum_{i=1}^l D_{1/4,i}^+(f_{1/4}) \beta_{1/4,i}^+(g_{\log T} x, 1) + \sum_{i=1}^l (D_{1/4,i}^-(f_{1/4}) - \frac{\log T}{2} D_{1/4,i}^+(f_{1/4})) \beta_{1/4,i}^-(g_{\log T} x, 1) \right| \leq C_r \|f\|_r T^{-\gamma}.$$

We then apply the above asymptotic formulas to the functions $f = \tau - \tau \circ g_s$. Since the action of the geodesic flow $g_{\mathbb{R}}$ preserves the splitting of Sobolev spaces into irreducible components, hence in particular, for all $s \in \mathbb{R}$, we have

$$f_\mu = \tau_\mu - \tau_\mu \circ g_s, \quad \text{for all } \mu \in \operatorname{spec}(\square).$$

It remains to compute the values of the invariant distributions $D_\mu^\pm \in \mathcal{D}'(H_\mu)$ on the function $\tau - \tau \circ g_s$ for any irreducible component H_μ of Casimir parameter $\mu > 1/4$. For $\mu \neq 1/4$, it follows from [12], Theorem 3.2 and Lemma 3.5 (see also [6], formulas (5) and (6)) that we have

$$D_\mu^\pm(\tau \circ g_s) = [(g_s)_*(D_\mu^\pm)](\tau) = \exp(-\frac{1 \pm \sqrt{1-4\mu}}{2} s) D_\mu^\pm(\tau),$$

and, for $\mu = 1/4$, we have

$$\begin{aligned} D_{1/4}^+(\tau \circ g_s) &= [(g_s)_*(D_{1/4}^+)](\tau) = e^{-\frac{s}{2}} D_{1/4}^+(\tau), \\ D_{1/4}^-(\tau \circ g_s) &= [(g_s)_*(D_{1/4}^-)](\tau) = e^{-\frac{s}{2}} [D_{1/4}^-(\tau) - \frac{s}{2} D_{1/4}^+(\tau)]. \end{aligned}$$

Since for every $r > 0$ there exists a constant $C_r > 0$ such that

$$\|\tau \circ g_s - \tau\| \leq C_r s \|\tau\|_r,$$

we can conclude the argument as follows. If $\mu_\tau < 1/4$, we set

$$\begin{aligned}\phi_\tau(s) &:= \exp\left(-\frac{1-\sqrt{1-4\mu_\tau}}{2}s\right) - 1, \quad \text{for all } s \in \mathbb{R}, \\ \beta_\tau(x) &:= (1 \pm \sqrt{1-4\mu_\tau}) \operatorname{Re} \sum_{i=1}^k D_i^-(\tau) \beta_i^-(x), \quad \text{for all } x \in M,\end{aligned}$$

so that from formula (B.1) we derive that

$$\left| T^{-\frac{1+\nu_f}{2}} \int_0^T (\tau - \tau \circ g_s) \circ h_t(x) dt - \phi_\tau(s) \beta_\tau(g_{\log T} x) \right| \leq C_r s \|\tau\|_r T^{-\gamma}.$$

If $\mu_\tau > 1/4$, then we set $\mathbf{v} = (\frac{v_n}{2})$ and, for all $(s, \theta, x) \in \mathbb{R} \times \mathbb{T}^\infty \times M$, we have

$$\beta_\tau(s, \theta, x) = \operatorname{Re} \sum_{n \in \mathbb{N}} \left[\exp\left(-\frac{1-\sqrt{1-4\mu_n}}{2}s\right) - 1 \right] D_{\mu_n}^+(\tau) \exp(i \frac{v_n}{2} \theta) \beta_{\mu_n}^+(x, 1)$$

so that, from formula (B.2), we derive that

$$\left| T^{-\frac{1}{2}} \int_0^T (\tau - \tau \circ g_s) \circ h_t(x) dt - \beta_\tau(s, \mathbf{v} \log T, g_{\log T} x) \right| \leq C_r s \|\tau\|_r T^{-\gamma}.$$

If $\mu_\tau = 1/4$, then we also set

$$\beta_\tau^{(1/4)}(x) := \sum_{i=1}^l D_{1/4,i}^+(\tau) \beta_{1/4,i}^+(x, 1), \quad \text{for all } x \in M,$$

so that, by formulas (B.2) and (B.3), we derive

$$\begin{aligned} \left| T^{-\frac{1}{2}} \int_0^T (\tau - \tau \circ g_s) \circ h_t(x) dt - \beta_\tau(s, \mathbf{v} \log T, g_{\log T} x) \right. \\ \left. - (e^{-s/2} - 1) \beta_\tau^{(1/4)}(g_{\log T} x) \log T \right| \leq C_r s \|\tau\|_r T^{-\gamma}. \end{aligned}$$

□

APPENDIX C. STRONG MOMO AND USIC PROPERTIES

We now consider a bounded arithmetic functions $\mathbf{u} : \mathbb{N} \rightarrow \mathbb{C}$ and a topological dynamical system (X, T) (i.e. X is a compact metric space and T is a homeomorphism of X).

The following notion has been introduced in [2]. MOMO is an acronym for Möbius Orthogonality of Moving Orbits.

Definition C.1 (strong MOMO). (X, T) satisfies the *strong \mathbf{u} -MOMO property* if for all $(b_k) \subset \mathbb{N}$ with $b_{k+1} - b_k \rightarrow \infty$, $(x_k) \subset X$ and $f \in C(X)$, we have

$$\lim_{K \rightarrow \infty} \frac{1}{b_K} \sum_{k < K} \left| \sum_{b_k \leq n < b_{k+1}} f(T^n x_k) \mathbf{u}(n) \right| = 0.$$

Proposition C.1. (X, T) satisfies strong \mathbf{u} -MOMO property if and only if it satisfies strong \mathbf{u} -USIC property.

Proof. To prove one implication, suppose that (X, T) does not satisfy \mathbf{u} -USIC. Then according to (9.4) there are: $\varepsilon_0 > 0$, $(M_\ell), (H_\ell)$ tending to ∞ with $H_\ell/M_\ell \rightarrow 0$, $f \in C(X)$ such that for a certain subsequence (ℓ_k) which we still denote by (ℓ) , we can find $x_\ell \in X$ such that

$$(C.1) \quad \frac{1}{M_\ell} \sum_{M_\ell \leq m < 2M_\ell} \left| \frac{1}{H_\ell} \sum_{m \leq h < m+H_\ell} f(T^h x_\ell) \mathbf{u}(h) \right| \geq \varepsilon_0.$$

We now proceed as in the proof of Theorem 5 in [1]. We have

$$\frac{1}{H_\ell} \sum_{r=0}^{H_\ell-1} \frac{1}{M_\ell/H_\ell} \sum_{M_\ell \leq m < 2M_\ell, m=r \bmod H_\ell} \left| \frac{1}{H_\ell} \sum_{m \leq h < m+H_\ell} f(T^h x_\ell) \mathbf{u}(h) \right| \geq \varepsilon_0.$$

Hence, for each ℓ , we can choose $0 \leq r_\ell < H_\ell$ such that

$$(C.2) \quad \frac{1}{M_\ell/H_\ell} \sum_{M_\ell \leq m < 2M_\ell, m=r_\ell \bmod H_\ell} \left| \frac{1}{H_\ell} \sum_{m \leq h < m+H_\ell} f(T^h x_\ell) \mathbf{u}(h) \right| \geq \varepsilon_0.$$

By passing to a subsequence if necessary, we can assume that $M_{\ell+1} > 2M_\ell + H_\ell$. The sequence (b_k) is defined as

$$\{b_1 < b_2 < \dots\} := \bigcup_{\ell} \{M_\ell \leq m < 2M_\ell + H_\ell, m = r_\ell \bmod H_\ell\}$$

(note that $b_{k+1} - b_k \geq H_\ell \rightarrow \infty$). Moreover, for all k such that $b_k \in [M_\ell, 2M_\ell + H_\ell]$, we set $x_k = x_\ell$. We let K_ℓ be the largest k so that $b_k < 2M_\ell$. Now, $b_{K_\ell}/(2M_\ell) \rightarrow 1$ (as $H_\ell/M_\ell \rightarrow 0$) and using (C.2), we obtain

$$\liminf_{\ell \rightarrow \infty} \frac{1}{b_{K_\ell}} \sum_{k < K_\ell} \left| \sum_{b_k \leq n < b_{k+1}} f(T^n x_k) \mathbf{u}(n) \right| \geq$$

$$\liminf_{\ell \rightarrow \infty} \frac{1}{2M_\ell} \sum_{M_\ell \leq m < 2M_\ell, m=r_\ell \bmod H_\ell} \left| \frac{1}{H_\ell} \sum_{m \leq h < m+H_\ell} f(T^h x_\ell) \mathbf{u}(h) \right| \geq \varepsilon_0/2.$$

We omit what happens between b_{K_ℓ} and $b_{K_\ell+1}$ as we are interested in an estimate from below.

To prove the other implication, we will need a lemma which has been shown in [2] (see Lemma 24 and its proof therein):

Lemma C.1. *Let $f \in C(X)$, $(x_k) \subset X$, $(b_k) \subset \mathbb{N}$, $b_{k+1} - b_k \rightarrow \infty$. Assume that*

$$\limsup_{K \rightarrow \infty} \frac{1}{b_{K+1}} \sum_{k \leq K} \left| \sum_{b_k \leq n < b_{k+1}} f(T^n x_k) \mathbf{u}(n) \right| > 0.$$

Then there exist $\delta_0 > 0$ and a subsequence (k_ℓ) such that the upper density

$$\eta := \bar{d} \left(\bigcup_{\ell \geq 1} [b_{k_\ell}, b_{k_\ell+1}) \right) > 0$$

and for each $\ell \geq 1$ we have

$$(C.3) \quad \frac{1}{b_{k_\ell+1} - b_{k_\ell}} \left| \sum_{n=b_{k_\ell}}^{b_{k_\ell+1}-1} f(T^n x_{k_\ell}) \mathbf{u}(n) \right| \geq \delta_0.$$

We now proceed, as before, by contradiction. That is, we suppose that the assumption of Lemma C.1 is satisfied. We need to select (M_ℓ) and (H_ℓ) . For that we use the assertion of that lemma. First of all (M_ℓ) will be chosen so that the upper density $\eta > 0$ of the set $\bigcup_{r \geq 1} [b_{k_r}, b_{k_r+1})$ is “realized”, that is, $|[1, M_\ell] \cap \bigcup_{r \geq 1} [b_{k_r}, b_{k_r+1})| \geq \frac{\eta}{2} M_\ell$ for each $\ell \geq 1$. Moreover, due to the assumption $b_{k+1} - b_k \rightarrow \infty$, we can assume (by passing to a subsequence of (M_ℓ) if necessary) that the union of intervals $[b_{k_r}, b_{k_r+1})$ whose length is short, say $\leq \ell$, has cardinality smaller than or equal to M_ℓ/ℓ . Finally, set $H_\ell = \sqrt{\ell}$.

Clearly $H_\ell/M_\ell \rightarrow 0$, and according to the strong \mathbf{u} -USIC property, most of the m in the interval $[1, M_\ell]$ is “good” in the sense that the relevant sums along $[m, m + H_\ell)$ are small, i.e., given $\varepsilon > 0$, for ℓ large enough, we have

$$\frac{1}{H_\ell} \sup_{x \in X} \left| \sum_{m \leq h < m + H_\ell} f(T^h x) \mathbf{u}(h) \right| < \varepsilon$$

for a $(1 - \varepsilon)$ -proportion of $m \in [1, M_\ell]$. We can make such “good” intervals of length H_ℓ disjoint by considering $m = s \bmod H_\ell$ and then choosing $0 \leq s_\ell < H_\ell$ so that most of the intervals $[m, m + H_\ell)$ with $m = s_\ell \bmod H_\ell$ are “good” (as in the proof of the necessity). Now, such “good” intervals will cover (with a small error) many of our “bad”, i.e. satisfying (C.3), intervals $[b_{k_r}, b_{k_r+1})$ (as the (upper) density of the union of such intervals is fixed, equal to η). However, if $[m, m + H_\ell)$ is good then $|\sum_{m \leq h < m + H_\ell} f(T^h y) \mathbf{u}(h)| = o(H_\ell)$ (with “o” which does not depend on m) regardless $y \in X$. It follows that we must have $|\sum_{b_{k_r} \leq n < b_{k_r+1}} f(T^n x_{k_r}) \mathbf{u}(n)| = o(b_{k_r+1} - b_{k_r})$, a contradiction. \square

Acknowledgements. The original motivation to start this project came from a question that Marina Ratner asked during the conference “*Homogeneous Dynamics, Unipotent Flows, and Applications*” held in October 2013 at the IAS in Jerusalem (and appeared as Question 7 in the survey [11]). We are thankful to her for asking this question in particular, and for her seminal and inspirational work more generally. We are also grateful to the IAS and to the conference organizers for making the event possible.

We are indebted to Giovanni Forni for explaining to us Lemma B.1 and how to deduce it from [6]. We would also like to thank him and Livio Flaminio for useful discussions.

M.L. is supported by an NCN grant. C.U. is supported by the ERC Starting Grant ChaParDyn, as well as by the Leverhulme Trust through a Leverhulme Prize and by the Royal Society through a Wolfson Research Merit Award. The research leading to these results has received funding from the European Research Council under the European Union Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 335989.

REFERENCES

- [1] E. H. EL ABDALAOUI, M. LEMAŃCZYK, AND T. DE LA RUE, *Automorphisms with Quasi-discrete Spectrum, Multiplicative Functions and Average Orthogonality Along Short Intervals*, Int. Math. Res. Notices 14 (2017), 4350–4368.
- [2] E. H. EL ABDALAOUI, J. KUŁAGA-PRZYMUS, M. LEMAŃCZYK, AND T. DE LA RUE, *Möbius disjointness for models of an ergodic system and beyond*, Israel J. Math., published online (2018), arXiv:1704.03506.
- [3] V. I. ARNOLD, *Topological and ergodic properties of closed 1-forms with incommensurable periods.*, Funktsional’nyi Analiz i Ego Prilozheniya, 25 (1991), pp. 1–12. (Translated in: *Functional Analyses and its Applications*, 25:2:81–90, 1991).
- [4] P. BERK, A. KANIGOWSKI, *Self-similarity problem for some surface flows*, preprint.
- [5] J. BOURGAIN, P. SARNAK, AND T. ZIEGLER, *Disjointness of Möbius from horocycle flows*, From Fourier analysis and number theory to Radon transforms and geometry, Dev. Math., vol. 28, Springer, New York, 2013, pp. 67–83.
- [6] A. BUFETOV, G. FORNI, *Limit Theorems for Horocycle Flows*, Annales scientifiques de l’ENS 47 (2014), 851–903.
- [7] I. P. CORNFELD, S. V. FOMIN, AND Y. G. SINAI, *Ergodic Theory*, Springer-Verlag, 1980.
- [8] C. DONG, A. KANIGOWSKI, *Rigidity of von Neumann special flows with one discontinuity*, preprint.
- [9] T. DOWNAROWICZ, J. SERAFIN, *Almost full entropy subshifts uncorrelated to the Möbius function*, to appear in Int. Math. Res. Notices, <https://arxiv.org/abs/1611.02084>.
- [10] B. FAYAD AND A. KANIGOWSKI, *Multiple mixing for a class of conservative surface flows*, Inventiones Mathematicae, (2015), pp. 1–60.
- [11] S. FERENCZI, J. KUŁAGA-PRZYMUS, M. LEMAŃCZYK, *Sarnak’s conjecture: what’s new*, in: Ergodic Theory and Dynamical Systems in their Interactions with Arithmetics and Combinatorics, CIRM

- Jean-Morlet Chair, Fall 2016, Editors: S. Ferenczi, J. Kułaga-Przymus, M. Lemańczyk, Lecture Notes in Mathematics 2213, Springer International Publishing, pp. 418.
- [12] L. FLAMINIO, G. FORNI, *Invariant distributions and time averages for horocycle flows*, Duke Math. J. 119, 2003, no. 3, 465–526.
 - [13] L. FLAMINIO, G. FORNI, *Orthogonal powers and Möbius conjecture for smooth time-changes of horocycle flows*, preprint.
 - [14] G. FORNI, A. KANIGOWSKI, *Multiple mixing and rigidity for reparametrizations of Heisenberg nilflows*, preprint.
 - [15] K. FRĄCZEK AND M. LEMAŃCZYK, *On mild mixing of special flows over irrational rotations under piecewise smooth functions*, Ergodic Theory and Dynamical Systems, 26 (2006), pp. 719–738.
 - [16] K. FRĄCZEK AND M. LEMAŃCZYK, *Ratner’s property and mild mixing for special flows over two-dimensional rotations*, J. Mod. Dyn., 4 (2010), pp. 609–635.
 - [17] H. FURSTENBERG, *Recurrence in ergodic theory and combinatorial number theory*, Princeton University Press, Princeton, N.J., 1981, M. B. Porter Lectures.
 - [18] P. GABRIEL, M. LEMAŃCZYK, K. SCHMIDT, *Extensions of cocycles for hyperfinite actions and applications*, Monatshefte Math. 123 (1997), 209–228.
 - [19] E. GLASNER, *Ergodic Theory via Joinings*, vol. 101 of Mathematical Surveys and Monographs, American Mathematical Society, Providence, RI, 2003.
 - [20] A. KANIGOWSKI, J. KUŁAGA-PRZYMUS, *Ratner’s property and mild mixing for smooth flows on surfaces*, Ergodic Theory Dynam. Systems (2015), FirstView.
 - [21] A. KANIGOWSKI, J. KUŁAGA-PRZYMUS, C. ULCIGRAI, *Multiple mixing and parabolic divergence in smooth area-preserving flows on higher genus surfaces*, to appear in JEMS.
 - [22] I. KATÁI, *A remark on a theorem of H. Daboussi*, Acta Math. Hungar., 47 (1986), 223–225.
 - [23] Y. KHINTCHIN, *Continued Fractions*, Chicago Univ. Press 1960
 - [24] ANDREY V. KOCHERGIN. Non-degenerate fixed points and mixing in flows on a 2-torus. Matem. aticheskii Sbornik, 194(8):83–112, 2003. (Translated in: Sb. Math., 194(8):1195–1224).
 - [25] L. KUIPERS, H. NIEDERREITEN, *Uniform distribution of sequences*, Wiley, London 1975.
 - [26] K. MATOMÄKI AND M. RADZIWIŁŁ, *Multiplicative functions in short intervals*, Ann. of Math. (2) **183** (2016), no. 3, 1015–1056.
 - [27] K. MATOMÄKI, M. RADZIWIŁŁ, AND T. TAO, *An averaged form of Chowla’s conjecture*, Algebra & Number Theory **9** (2015), 2167–2196.
 - [28] B. MARCUS, *The horocycle flow is mixing of all degrees*, Invent. Math. **46** (1978), no. 3, 201–209.
 - [29] S. P. NOVIKOV, *The Hamiltonian formalism and a multivalued analogue of Morse theory*, (Russian) Uspekhi Matematicheskikh Nauk, 37 (1982), pp. 3–49. (Translated in: *Russian Mathematical Surveys*, 37 No 5:1–56, 1982).
 - [30] M. RATNER, *Horocycle flows, joinings and rigidity of products*, Ann. of Math. (2), 118 (1983), pp. 277–313.
 - [31] M. RATNER, *Rigid reparametrizations and cohomology for horocycle flows*, Invent. Math., Vol. 88, (2) 1987, pp. 341–374.
 - [32] D. RAVOTTI, *Quantitative mixing for locally hamiltonian flows with saddle loops on compact surfaces*. Preprint arXiv:1610.08743, 2016.
 - [33] V. V. RYZHIKOV AND J.-P. THOUVENOT, *Disjointness, divisibility, and quasi-simplicity of measure-preserving actions*, Funktsional. Anal. i Prilozhen., 40 (2006), 85–89.
 - [34] P. SARNAK, *Three lectures on the Möbius function, randomness and dynamics*, <http://publications.ias.edu/sarnak/>.
 - [35] Y. G. SINAI AND K. M. KHANIN, *Mixing for some classes of special flows over rotations of the circle.*, Funktsional’nyi Analiz i Ego Prilozheniya, 26 (1992), pp. 1–21. (Translated in: *Functional Analysis and its Applications*, 26:3:155–169, 1992).
 - [36] J.-P. THOUVENOT, *Some properties and applications of joinings in ergodic theory*, in Ergodic theory and its connections with harmonic analysis (Alexandria, 1993), vol. 205 of London Math. Soc. Lecture Note Ser., Cambridge Univ. Press, Cambridge, 1995, pp. 207–235.
 - [37] C. ULCIGRAI, *Mixing of asymmetric logarithmic suspension flows over interval exchange transformations*, Ergodic Theory Dynam. Systems, 27 (2007), pp. 991–1035.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MARYLAND AT COLLEGE PARK, COLLEGE PARK,
MD 20740, USA

E-mail address: `adkanigowski@gmail.com`

FACULTY OF MATHEMATICS AND COMPUTER SCIENCE, NICOLAUS COPERNIC UNIVERSITY UL. CHOP-
INA 12/18 87-100 TORUN, POLAND

E-mail address: `mlem@mat.umk.pl`

INSTITUT FÜR MATHEMATIK, UNIVERSITÄT ZÜRICH WINTERTHURERSTRASSE 190, CH-8057 ZÜRICH

SCHOOL OF MATHEMATICS, UNIVERSITY OF BRISTOL, HOWARD HOUSE, QUEENS AVE BS8 1SD
BRISTOL, UK

E-mail address: `corinna.ulcigrai@gmail.com`